

F:2202-2209/Region: nucleotide-binding motif A (P-loop)  
F:2530-2537/Region: nucleotide-binding motif A (P-loop)

Query Match 56.7%; Score 38; DB 2; Length 4568;  
Best Local Similarity 41.7%; Pred. No. 2.7e+02;  
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWXRNMKVR 12  
|||  
:|:|:  
Db 1852 CFQWQSLRYIQ 1863

## RESULT 10

B60950  
apolipoprotein B-100 - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 31-Dec-1993 #sequence\_revision 09-Sep-1994 #text\_change 07-Oct-1994  
C:Accession: B60950  
R:Law, A.; Scott, J.  
J. Lipid Res. 31, 1109-1120, 1990  
A:Title: A cross-species comparison of the apolipoprotein B domain that binds to the LDL  
A:Reference number: A60950; MUID:90324804; PMID:2373961  
A:Accession: B60950  
A:Molecule type: mRNA  
A:Residues: 1-274 <LAW>  
A>Note: authors translated the codon ATA for residue 8 as Val  
C:Superfamily: apolipoprotein B  
C:Keywords: atherosclerosis; calcium; cholesterol metabolism; chylomicron; glycoprotein;

Query Match 55.2%; Score 37; DB 2; Length 274;  
Best Local Similarity 66.7%; Pred. No. 27;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 WXRNMKVR 12  
|||  
:|:|:  
Db 39 WDRNLKFR 47

## RESULT 11

C70655  
probable monooxygenase - Mycobacterium tuberculosis (strain H37Rv)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C:Accession: C70655  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: C70655  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-489 <COL>  
A:Cross-references: GB:Z83864; GB:AL123456; MUID:g3261687; PIDN:CAB06212.1; PID:e301250;  
A:Experimental source: strain H37Rv  
C:Genetics:  
A:Gene: RV3854c

Query Match 55.2%; Score 37; DB 2; Length 489;  
Best Local Similarity 54.5%; Pred. No. 47;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWXRNMKVR 11  
|||  
:|:|:  
Db 253 CQWPRRMKRM 263

## RESULT 12

T39801  
hypothetical sh3-containing protein - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T39801  
R:McDougal, R.C.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.  
submitted to the EMBL Data Library, July 1999  
A:Reference number: 221880  
A:Accession: T39801  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-501 <MCD>  
A:Cross-references: EMBL:AL109731; PIDN:CAB52037.1; GSPDB:GN00067; SPDB:SPBC19C2.10  
A:Experimental source: strain 972h-; cosmid c19C2  
C:Genetics:  
A:Gene: SPDB:SPBC19C2.10  
A:Map position: 2  
A:Introns: 196/3

Query Match 55.2%; Score 37; DB 2; Length 501;  
Best Local Similarity 60.0%; Pred. No. 48;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWXRNMK 10  
|||  
:|:|:  
Db 50 CTRWYRNDK 59

## RESULT 13

C84325  
hypothetical protein Vng1732c [imported] - Halobacterium sp. NRC-1  
C:Species: Halobacterium sp. NRC-1  
C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: C84325  
R:Ng, W.W.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, J.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl, Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; L  
A:Title: Genome sequence of Halobacterium species NRC-1.  
A:Reference number: A84160; MUID:20504483; PMID:11016950  
A:Accession: C84325  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-584 <STO>  
A:Cross-references: GB:AE004437; MUID:g10581192; PIDN:AAG19967.1; GSPDB:GN00138  
C:Genetics:  
A:Gene: VNG1732C

Query Match 55.2%; Score 37; DB 2; Length 584;  
Best Local Similarity 41.7%; Pred. No. 56;  
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWXRNMKVR 12  
|||  
:|:|:  
Db 445 CFTWRDMKVR 456

## RESULT 14

C96582  
F1511.22 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: C96582  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; anson, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: C96582

```
;Status: preliminary
;Molecule type: DNA
;Residues: 1-124 <STO>
;Cross-references: GB:AE005173; NID:g4587555; PIDN:AAD25786.1; GSPDB:GN00141
;Genetics:
;Gene: F15I1.22
;Map position: 1

Query Match      53.7%; Score 36; DB 2; Length 124;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Y      2 FQWYRNVRK 11
      |||
b     11 FWRSRRKI 20

RESULT 15
D2346
;Hypothetical protein alr4323 [imported] - Nostoc sp. (strain PCC 7120)
;Species: Nostoc sp.
;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
;Accession: AD2346
;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
NA Res. 8, 205-213, 2001
;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
;Reference number: AB1807; MUID:21595285; PMID:11759840
;Accession: AD2346
;Status: preliminary
;Molecule type: DNA
;Residues: 1-298 <KUR>
;Cross-references: GB:BA000019; PIDN:BAE76022.1; PID:gi17133459; GSPDB:GN00179
;Experimental source: strain PCC 7120
;Genetics:
;Gene: alr4323

Query Match      53.7%; Score 36; DB 2; Length 298;
Best Local Similarity 66.7%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Y      2 FQWYRNVRK 10
      |||
b     163 FHWQRYRK 171

earch completed: February 21, 2003, 08:02:44
ob time : 11.6047 secs
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

DM protein - protein search, using sw model

Run on: February 21, 2003, 07:28:06 ; Search time 5.2093 Seconds  
(without alignments)  
95.544 Million cell updates/sec

Title: US-09-743-107B-88

Perfect score: 67

Sequence: 1 CFQXRNRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	65	97.0	711	1 TRFL_HUMAN	P02788 homo sapien
2	48	71.6	708	1 TRFL_CAMDR	Q9tumo camelus dro
3	48	71.6	708	1 TRFL_CAPHI	Q29477 capra hircu
4	40	59.7	292	1 NLA_DROME	Q9xz18 drosophila
5	40	59.7	695	1 TRFL_HORSE	Q77811 equus cabal
6	39	58.2	707	1 TRFL_MOUSE	P08071 mus musculu
7	38	56.7	146	1 RPOB_LIBAP	P41187 liberibacte
8	38	56.7	783	1 YNR2_CAEEL	Q21988 caenorhabdi
9	38	56.7	4568	1 DYHB_CHURE	Q39565 chlamydomon
10	36	53.7	365	1 IA34_HUMAN	P30453 homo sapien
11	36	53.7	369	1 SPI1_MYXVL	P12393 myxoma viru
12	36	53.7	428	1 SYH_CHLMU	Q9pj19 chlamydia m
13	36	53.7	749	1 VP4_ROTGA	Q04916 rotavirus (
14	35	52.2	62	1 RL2B_THETN	Q8rsul thermanser
15	35	52.2	966	1 VIA_EBWV	Q00020 broad bean
16	35	52.2	1135	1 PHYC_SORBI	P93528 sorghum bic
17	34	50.7	214	1 VIF_SIVS4	P12505 simian immu
18	34	50.7	422	1 PAF4_CHICK	Q90678 g platelet-
19	34	50.7	428	1 SYH_CHLTPN	Q84547 chlamydia t
20	34	50.7	430	1 SYH_CHLTPN	Q92701 chlamydia p
21	34	50.7	455	1 YKYL_CAEEL	Q19310 caenorhabdi
22	34	50.7	480	1 YQSL_CAEEL	Q09309 caenorhabdi
23	34	50.7	500	1 TLCE_RICPR	O05962 rickettsia
24	34	50.7	522	1 COR1_SCHPO	O13688 schistosacch
25	34	50.7	663	1 PD11_HUMAN	Q9ulc6 homo sapien
26	34	50.7	765	1 Y008_HUMAN	Q15398 homo sapien
27	34	50.7	989	1 TL00_HUMAN	Q75448 homo sapien
28	34	50.7	1202	1 JAG2_RAT	P97607 rattus norv
29	34	50.7	1238	1 JAG2_HUMAN	Q9y219 homo sapien
30	34	50.7	1247	1 JAG2_MOUSE	Q9gyes mus musculu
31	34	50.7	2151	1 RRPL_SEOUB	P27314 seoul virus
32	34	50.7	4568	1 DYHC_CAEEL	Q19020 caenorhabdi
33	33.5	50.0	727	1 KDGA_RAT	P51556 rattus norv

34 33.5 50.0 730 1 KDGA\_MOUSE O88673 mus musculu  
35 33 49.3 62 1 RL28\_STAAM Q99up4 staphylococ  
36 33 49.3 144 1 REV\_OMYVS P16903 ovine lenti  
37 33 49.3 172 1 Y708\_METJA Q58118 methanococ  
38 33 49.3 211 1 PDOI\_ECOLI P32174 escherichia  
39 33 49.3 224 1 GTXA\_ARATH P46421 arabidopsis  
40 33 49.3 238 1 CCDA\_BACHD Q9kd18 bacillus ha  
41 33 49.3 246 1 Y495\_SYNY3 Q55195 synecocyst  
42 33 49.3 275 1 VAL6\_VACCV P16710 vaccinia vi  
43 33 49.3 378 1 VAL6\_VACCC P20993 vaccinia vi  
44 33 49.3 573 1 UREL\_LACFE P28929 lactobacill  
45 33 49.3 648 1 Z202\_HUMAN O95125 homo sapien

#### ALIGNMENTS

RESULT 1  
TRFL\_HUMAN STANDARD; PRT: 711 AA.  
AC P02788; Q16780; Q16785; Q16786; Q16789; Q00756; Q9H1Z3; Q96KZ4;  
AC Q96KZ5;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DB Lactotransferrin precursor (Lactoferrin) [Contains: Lactoferrin A;  
DB Lactoferrin B; Lactoferrin C].  
GN LTF OR LF.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OK NCBI TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RX MEDLINE=90384839; PubMed=2402455;  
RA Rey M.W., Woloshuk S.L., de Boer H.A., Pieper F.R.;  
RT "Complete nucleotide sequence of human mammary gland lactoferrin.";  
RL Nucleic Acids Res. 18:5288-5288(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RA Cho Y.Y.;  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Connely O.M.;  
RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RA Liang Q., Jimenez-Flores R., Richardson T.;  
RN [5]  
RP "Molecular cloning and sequence analysis of human lactoferrin.";  
RL Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Bone marrow;  
RA Wei X., Han J., Rado T.A.;  
RN [7]  
RP "Human neutrophil lactoferrin coding and 5' flanking region DNA  
sequences";  
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.  
RN [8]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RA Cheng H., Chen X., Huan L.;  
RN [9]  
RP "cDNA cloning and sequence analysis of human lactoferrin.";  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [10]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Prostate;  
RA Strausberg R.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

RC SEQUENCE OF 3-711 FROM N.A.  
 RC TISSUE=Mammary Gland;  
 RA MEDLINE=90326549; PubMed=2374734;  
 RA Powell M.J., Oden J.E.;  
 RT "Nucleotide sequence of human lactoferrin cDNA.";  
 RL Nucleic Acids Res. 18:4013-4013(1990).  
 RN [9]  
 RN SEQUENCE OF 20-711.  
 RP MEDLINE=8507667; PubMed=6510420;  
 RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Schoentgen F.,  
 RA Legrand D., Spik G., Montreuil J., Jolles P.;  
 RA "Human lactotransferrin: amino acid sequence and structural  
 RT comparisons with other transferrins.";  
 RL Eur. J. Biochem. 145:659-666(1984).  
 RN [10]  
 RP PRELIMINARY SEQUENCE OF 20-73; 134-171; 257-278 AND 347-530.  
 RA Metz-Boutigue M.-H., Mazurier J., Jolles J., Spik G., Montreuil J.,  
 RA Jolles P.;  
 RT "The present state of the human lactotransferrin sequence. Study and  
 RT alignment of the cyanogen bromide fragments and characterization of  
 RT N- and C-terminal domains.";  
 RL Biochim. Biophys. Acta 670:243-254(1981).  
 RN [11]  
 RP SEQUENCE OF 609-711.  
 RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,  
 RA Jolles P.;  
 RT "An 88 amino acid long C-terminal sequence of human  
 RT lactotransferrin.";  
 RL FEBS Lett. 142:107-110(1982).  
 RN [12]  
 RP SEQUENCE OF 436-711 FROM N.A.  
 RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,  
 RA Jolles P.;  
 RT "Isolation of lactoferrin cDNA from a human myeloid library and  
 RT expression of mRNA during normal and leukemic myelopoiesis.";  
 RL Blood 70:989-993(1987).  
 RN [13]  
 RP SEQUENCE OF 237-711 FROM N.A.  
 RA McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,  
 RA Nhan M., Parnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,  
 RA Gnoj L., la Bastide M., Kaplan N., Greco T., Touchman J., Mazny D.,  
 RA Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M., Porcel B.M., U.,  
 RA Dragan Y., Giacalone J., Pae A., Powell E., Solinsky K.A., Desliva U.,  
 RA Diaz-Perez S., Zhou X., Yu Y., Watanabe M., Doggett N., Garcia D.,  
 RA Sagripanti J.L.;  
 RA Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [14]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND REVISIONS.  
 RA MEDLINE=90084528; PubMed=2585506;  
 RA Anderson B.F., Baker H.M., Norris G.E., Rice D.W., Baker E.N.;  
 RT "Structure of human lactoferrin: crystallographic structure analysis  
 RT and refinement at 2.8-A resolution.";  
 RL J. Mol. Biol. 209:711-734(1989).  
 RN [15]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RA Haridas M., Anderson B.F., Baker E.N.;  
 RT "Structure of human diferric lactoferrin refined at 2.2-A  
 RT resolution.";  
 RL Acta Crystallogr. D 51:629-646(1995).  
 RN [16]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 20-353.  
 RA MEDLINE=97156796; PubMed=9003186;  
 RA Nicholson H., Anderson B.F., Bland T., Shewry S.C., Tweedie J.W.,  
 RA Baker E.N.;  
 RT "Mutagenesis of the histidine ligand in human lactoferrin: iron  
 RT binding properties and crystal structure of the histidine-  
 RT 253-->methionine mutant.";  
 RL Biochemistry 36:341-346(1997).  
 RN [17]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RA MEDLINE=99190892; PubMed=10089347;  
 RX Sun X.L., Baker H.M., Shewry S.C., Jameson G.B., Baker E.N.;  
 RT "Structure of recombinant human lactoferrin expressed in Aspergillus  
 RT awamori.";  
 RL Acta Crystallogr. D 55:403-407(1999).  
 RN [18]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 RA MEDLINE=99192677; PubMed=10089508;  
 RA Jameson G.B., Anderson B.F., Norris G.E., Thomas D.H., Baker E.N.;  
 RT "Structure of human apolactoferrin at 2.0-A resolution. Refinement  
 RT and analysis of ligand-induced conformational change.";  
 RL Acta Crystallogr. D 54:1319-1335(1998).  
 RN [19]  
 RP CHARACTERIZATION OF LACTOFERROXINS.  
 RX MEDLINE=91166929; PubMed=1369293;  
 RA Tani F., Iio K., Chiba H., Yoshikawa M.;  
 RT "Isolation and characterization of opioid antagonist peptides derived  
 RT from human lactoferrin.";  
 RL Agric. Biol. Chem. 54:1803-1810(1990).  
 RN [20]  
 RP VARIANTS THR-30 AND ARG-48.  
 RX PubMed=9873069;  
 RA Klintworth G.K., Sommer J.R., Obrian G., Han L., Ahmed M.N.,  
 RA Qumsiyeh M.B., Lin P.-Y., Baeti S., Reddy M.K., Kanai A., Hotta Y.,  
 RA Sugar J., Kumaramanickavel G., Munier P., Schorderet D.F.,  
 RA El Matiri L., Iwata F., Kaiser-Kupfer M., Nagata M., Nakayasu K.,  
 RA Hejtmancik J.P., Teng C.T.;  
 RT "Familial subepithelial corneal amyloidosis (gelatinous drop-like  
 RT corneal dystrophy): exclusion of linkage to lactoferrin gene.";  
 RL Mol. Vision 4:31-32(1998).  
 RN [21]  
 RP FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY BICARBONATE.  
 CC [22]  
 RP FUNCTION: LACTOFERROXINS A, B AND C HAVE OPIOID ANTAGONIST  
 CC ACTIVITY. LACTOFERROXIN A SHOWS PREFERENCE FOR MU-RECEPTORS, WHILE  
 CC LACTOFERROXIN B AND C HAVE SOMEWHAT HIGHER DEGREES OF PREFERENCE  
 CC FOR KAPPA-RECEPTORS THAN FOR MU-RECEPTORS.  
 CC [23]  
 RP SUBUNIT: MONOMER.  
 CC [24]  
 RP SUBCELLULAR LOCATION: Secreted.  
 CC [25]  
 RP DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC [26]  
 RP SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
 CC [27]  
 RP This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL cutation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC [28]  
 RP EMBL; X53961; CAA37914.1; -  
 CC EMBL; U07643; AAB60324.1; -  
 DR EMBL; M33150; AAB36159.1; -  
 DR EMBL; M83202; AAA59511.1; -  
 CC EMBL; M83205; AAA58656.1; -  
 DR EMBL; M18642; AAA86665.1; -  
 CC EMBL; AF332168; AAG48753.1; -  
 DR EMBL; BC015822; AAH15822.1; -  
 DR EMBL; BC015823; AAH15823.1; -  
 CC EMBL; M73700; AAA59479.1; -  
 DR EMBL; X52941; CAA37116.1; -  
 CC EMBL; U95626; AAB57795.1; -  
 DR PIR; S11228; TFHUL.  
 DR PDB; 1LCF; 31-AUG-94.  
 DR PDB; 1LCT; 31-OCT-93.  
 DR PDB; 1LFG; 31-JUL-94.  
 DR PDB; 1LFH; 31-OCT-93.  
 DR PDB; 1LFI; 31-OCT-93.  
 DR PDB; 1LGB; 31-AUG-94.  
 DR PDB; 1LGC; 31-AUG-94.  
 DR PDB; 1BKA; 08-NOV-96.  
 DR PDB; 1DSN; 08-MAR-96.  
 DR PDB; 1HSE; 12-MAR-97.  
 DR PDB; 1VFD; 21-APR-97.

FT	DISULFID	377	390		BY SIMILARITY.
FT	DISULFID	424	703		BY SIMILARITY.
FT	DISULFID	444	666		BY SIMILARITY.
FT	DISULFID	476	551		BY SIMILARITY.
FT	DISULFID	500	694		BY SIMILARITY.
FT	DISULFID	510	524		BY SIMILARITY.
FT	DISULFID	521	534		BY SIMILARITY.
FT	DISULFID	592	606		BY SIMILARITY.
FT	DISULFID	644	649		BY SIMILARITY.
FT	METAL	79	79		IRON 1 (BY SIMILARITY).
FT	METAL	111	111		IRON 1 (BY SIMILARITY).
FT	METAL	211	211		IRON 1 (BY SIMILARITY).
FT	METAL	272	272		IRON 1 (BY SIMILARITY).
FT	METAL	414	414		IRON 2 (BY SIMILARITY).
FT	METAL	452	452		IRON 2 (BY SIMILARITY).
FT	METAL	545	545		IRON 2 (BY SIMILARITY).
FT	METAL	614	614		IRON 2 (BY SIMILARITY).
FT	BINDING	140	140		ANION (BY SIMILARITY).
FT	BINDING	482	482		ANION (BY SIMILARITY).
FT	CARBOHYD	252	252		N-LINKED (GLCNAC. .)
FT	CARBOHYD	385	385		N-LINKED (GLCNAC. .)
FT	CARBOHYD	537	537		N-LINKED (GLCNAC. .)
FT	CARBOHYD	594	594		F -> S (IN REF. 2).
FT	CONFLICT	261	261		G -> A (IN REF. 2).
FT	CONFLICT	330	330		S -> P (IN REF. 2).
FT	CONFLICT	330	330		LIS -> PLF (IN REF. 2).
FT	CONFLICT	492	494		L -> P (IN REF. 2).
FT	CONFLICT	506	506		A -> P (IN REF. 2).
FT	CONFLICT	609	609		R -> Q (IN REF. 2).
FT	CONFLICT	642	642		
FT	CONFLICT	642	642		
FT	SEQUENCE	708 AA;	72711 MW;		0B0C175A0B69D430 CRC64;
QY	1 CFQWXRNRKVR	12			
Dd	38 CAQWRMRKVR	49			
RESULT 3					
ID	TRFL_CAPHI	STANDARD;	PRT;	708 AA.	
AD	Q29477; Q29479;				
DT	15-DEC-1998 (Rel. 37, Created)				
DT	15-DEC-1998 (Rel. 37, Last sequence update)				
DE	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Lactotransferrin precursor (Lactoferrin).				
GN	Ltf.				
OS	Capra hircus (Goat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelestomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoides;				
OX	Bovidae; Caprinae; Capra.				
RN	NCBI_TaxID=9925;				
RP	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Mammary gland;				
RA	Lee T., Yu S., Kim S., Lee K., Yu D.;				
RL	Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.				
RL	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Mammary gland;				
RX	MEDLINE=94380047; PubMed=8093048;				
RT	le Provost F., Nocart M., Guerin G., Martin P.;				
RT	"Characterization of the goat lactoferrin cDNA. Assignment of the				
RT	relevant locus to bovine U12 syntenic group.";				
RL	Biochem. Biophys. Res. Commun. 203:1324-1332(1994).				
CC	-!- FUNCTION: TRANSFERINS ARE IRON BINDING TRANSPORT PROTEINS WHICH				
CC	CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING				
CC	OF AN ANION, USUALLY BICARBONATE.				
CC	-!- SUBUNIT: MONOMER (BY SIMILARITY).				
CC	-!- SUBCELLULAR LOCATION: Secreted.				



C or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

R EMBL; AF147700; AAD33987.1; -;  
R EMBL; AE003712; AAF55285.1; -;  
R FlyBase; FBgn0026629; nla.  
Q SEQUENCE 292 AA; 31423 MW; 64F1BBF5F6EA6CF9 CRC64;

Query Match 59.7%; Score 40; DB 1; Length 292;  
Best Local Similarity 54.5%; Pred. No. 2.6;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Y 2 FQWXRNMKVR 12  
||| | : : :  
b 150 FQWLRSPRLR 160

RESULT 5

D TRFL HORSE STANDARD; PRT; 695 AA.

C 07811;

T 15-JUN-1999 (Rel. 38, Created)

T 15-JUN-1999 (Rel. 38, Last sequence update)

T 15-JUN-2002 (Rel. 41, Last annotation update)

B Lactotransferrin precursor (Lactoferrin) (Fragment).

N LTP.

S Equus caballus (Horse).

C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

C Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

X NCBI\_TaxID=9796;

N [1]

P SEQUENCE FROM N.A.

A Paramasivam M., Srinivasan A., Yadav M.P., Singh T.P.;

T "CDNA sequence of mare lactoferrin";

L Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

N [2]

P X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).

C TISSUE=Milk;

X MEDLINE=99296631; PubMed=10366507;

U Sharma A.K., Paramasivam M., Srinivasan A., Yadav M.P., Singh T.P.;

T "Three-dimensional structure of mare diferric lactoferrin at 2.6-A

resolution.";

J. Mol. Biol. 289:303-317(1999).

X - FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH

CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING

OF AN ANION, USUALLY BICARBONATE.

X - SUBUNIT: MONOMER.

X - SUBCELLULAR LOCATION: Secreted.

X - DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.

X - SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AJ010930; CAA09407.1; -;

PDB; 1B1X; 02-DEC-98.

PDB; 1B7U; 02-FEB-99.

PDB; 1B7Z; 02-FEB-99.

InterPro; IPR001156; Transferrin.

Pfam; PF00405; transferrin; 2.

PRINTS; PR00422; TRANSFERRIN.

SMART; SMO0094; TR\_FER; 2.

PROSITE; PS00205; TRANSFERRIN\_1; 2.

PROSITE; PS00206; TRANSFERRIN\_2; 2.

PROSITE; PS00207; TRANSFERRIN\_3; 1.

Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;

Signal; 3d-structure.

NON\_TER 1

SIGNAL <1 6

RP

RA

FT CHAIN 7 695

FT REPEAT 7 350

FT REPEAT 351 695

FT DISULFID 15 51

FT DISULFID 25 42

FT DISULFID 121 204

FT DISULFID 163 179

FT DISULFID 166 189

FT DISULFID 176 187

FT DISULFID 237 251

FT DISULFID 354 386

FT DISULFID 364 377

FT DISULFID 411 690

FT DISULFID 431 653

FT DISULFID 463 536

FT DISULFID 487 681

FT DISULFID 497 511

FT DISULFID 508 521

FT DISULFID 579 593

FT DISULFID 631 636

FT METAL 66 66

FT METAL 98 98

FT METAL 198 198

FT METAL 259 259

FT METAL 401 401

FT METAL 439 439

FT METAL 532 532

FT METAL 601 601

FT BINDING 127 127

FT BINDING 469 469

FT CARBOHYD 143 143

FT CARBOHYD 287 287

FT CARBOHYD 482 482

SQ SEQUENCE 695 AA; 75991 MW; 07BB84D50E1B165D CRC64;

IRON 1 (BY SIMILARITY).

IRON 1 (BY SIMILARITY).

IRON 1 (BY SIMILARITY).

IRON 1 (BY SIMILARITY).

IRON 2 (BY SIMILARITY).

IRON 2 (BY SIMILARITY).

IRON 2 (BY SIMILARITY).

IRON 2 (BY SIMILARITY).

ANION (BY SIMILARITY).

ANION (BY SIMILARITY).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 59.7%; Score 40; DB 1; Length 695;

Best Local Similarity 58.3%; Pred. No. 6.1;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWXRNMKVR 12

DB 25 CAKFORNMKVR 36

RESULT 6

TRFL MOUSE STANDARD; PRT; 707 AA.

AC P08071; P70690; Q61799; Q922P2;

DT 01-AUG-1988 (Rel. 08, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Lactotransferrin precursor (Lactoferrin)

GN LTP.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OK NCBI\_TaxID=10090;

RP SEQUENCE FROM N.A.

RC TISSUE=Uterus;

RC MEDLINE=8728033; PubMed=3611056;

RA Pentecost B.T., Teng C.T.;

RT "Lactotransferrin is the major estrogen inducible protein of mouse

uterine secretions.";

RL J. Biol. Chem. 262:10134-10139(1987).

RL [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Uterus;

RA Moriishi K.;

RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.

RP SEQUENCE FROM N.A.

RA Strausberg R.;

Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 (4)  
 SEQUENCE OF 1-14 FROM N.A.  
 MEDLINES=92042099; PubMed=1939212;  
 Liu Y., Teng C.T.;  
 RT "Characterization of estrogen-responsive mouse lactoferrin promoter.";  
 J. Biol. Chem. 266:21880-21885(1991).  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 OF AN ANION, USUALLY BICARBONATE.  
 CC -!- SUBUNIT: MONOMER.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
 CC  
 EMBL: J03298; AAA40525.1; -  
 EMBL: D88510; BAA13633.1; -  
 EMBL: BC006904; AAH06904.1; -  
 EMBL: M74778; AAA39427.1; -  
 PIR: A28438; A28438.  
 HSP: P02788; ICB6.  
 MGD: MGI:98837; Lef.  
 InterPro: IPR001156; Transferrin.  
 Pfam: PF00405; transferrin; 2.  
 SMART: SM00094; TR FER; 2.  
 DR PROSITE; PS00205; TRANSFERRIN\_1; 1.  
 DR PROSITE; PS00206; TRANSFERRIN\_2; 2.  
 DR PROSITE; PS00207; TRANSFERRIN\_3; 2.  
 KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 KW Signal.  
 FT SIGNAL 1 19 BY SIMILARITY.  
 FT CHAIN 20 707 LACTOTRANSFERRIN.  
 FT REPEAT 20 357 1.  
 FT REPEAT 358 707 2.  
 FT DISULFID 27 63 BY SIMILARITY.  
 FT DISULFID 37 54 BY SIMILARITY.  
 FT DISULFID 133 216 BY SIMILARITY.  
 FT DISULFID 175 191 BY SIMILARITY.  
 FT DISULFID 188 199 BY SIMILARITY.  
 FT DISULFID 249 263 BY SIMILARITY.  
 FT DISULFID 366 398 BY SIMILARITY.  
 FT DISULFID 375 389 BY SIMILARITY.  
 FT DISULFID 423 702 BY SIMILARITY.  
 FT DISULFID 443 665 BY SIMILARITY.  
 FT DISULFID 475 550 BY SIMILARITY.  
 FT DISULFID 499 693 BY SIMILARITY.  
 FT DISULFID 509 523 BY SIMILARITY.  
 FT DISULFID 520 533 BY SIMILARITY.  
 FT DISULFID 591 605 BY SIMILARITY.  
 FT DISULFID 643 648 BY SIMILARITY.  
 FT METAL 78 78 IRON 1 (BY SIMILARITY).  
 FT METAL 110 110 IRON 1 (BY SIMILARITY).  
 FT METAL 210 210 IRON 1 (BY SIMILARITY).  
 FT METAL 271 271 IRON 1 (BY SIMILARITY).  
 FT METAL 413 413 IRON 2 (BY SIMILARITY).  
 FT METAL 451 451 IRON 2 (BY SIMILARITY).  
 FT METAL 544 544 IRON 2 (BY SIMILARITY).  
 FT METAL 613 613 IRON 2 (BY SIMILARITY).  
 FT BINDING 139 139 ANION (POTENTIAL).  
 FT BINDING 481 481 ANION (POTENTIAL).  
 FT CARBOHYD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 494 494 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 1 2 MR -> IOG (IN REF. 1).  
 FT CONFLICT 25 25 R -> Q (IN REF. 2).

FT CONFLICT 82 82 M -> L (IN REF. 2).  
 FT CONFLICT 359 359 S -> T (IN REF. 2).  
 FT CONFLICT 382 382 A -> D (IN REF. 1).  
 FT CONFLICT 449 449 E -> G (IN REF. 2).  
 FT CONFLICT 629 629 L -> V (IN REF. 1).  
 SQ SEQUENCE 707 AA; 77865 MW; F26AE0340A4C19A8 CRC64;  
 Query Match 58.2%; Score 39; DB 1; Length 707;  
 Best Local Similarity 54.5%; Pred. No. 9.6;  
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 CFQWXRNRKV 11  
 DB 37 CLRQWNRKV 47  
 RESULT 7  
 RPOB\_LIBAF  
 ID RPOB\_LIBAF STANDARD; PRT; 146 AA.  
 AC P41187;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE DNA-directed RNA polymerase beta chain (EC 2.7.7.6) (Transcriptase  
 beta chain) (RNA polymerase beta subunit) (Fragment).  
 DE RPOB.  
 OS Liberibacter africanus (Liberibacter africanus).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Candidatus Liberibacter.  
 OC NCBI\_TaxID=34020;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Nelspruit;  
 RA Planet P., Jaqueux S., Bove J.M., Garnier M.;  
 RT "Detection and characterization of the African Citrus Greening  
 RT Liberibacter by amplification, cloning and sequencing of the rplKJL-  
 RT rpoBC operon";  
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
 RC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION  
 OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS  
 SUBSTRATES.  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 (RNA) (N).  
 CC -!- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE  
 ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1  
 BETA' CHAIN.  
 CC -!- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
 CC  
 EMBL: U09675; AAA19857.1; -  
 DR InterPro: IPR001572; RNA pol B.  
 DR Pfam: PF00562; RNA pol B; 1.  
 DR PROSITE; PS01166; RNA\_POL\_BETA; PARTIAL.  
 KW Transferase; transcription; DNA-directed RNA polymerase.  
 FT NON\_TER 146 146  
 SQ SEQUENCE 146 AA; 16731 MW; 815D6EDD66FD8943 CRC64;  
 Query Match 56.7%; Score 38; DB 1; Length 146;  
 Best Local Similarity 60.0%; Pred. No. 3;  
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 CFQWXRNRKV 10  
 DB 10 CVQWNRGARK 19



```

RT with serology."
RL Tissue Antigens 41:72-80(1993).
CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -1- POLYMORPHISM: THE FOLLOWING ALLELES OF AW-34 ARE KNOWN: A*3401
CC (AW-34.1) AND A*3402 (AW34.2). THE SEQUENCE SHOWN IS THAT OF
CC A*3401.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X61704; CAA43873.1; -
DR EMBL; X61705; CAA43874.1; -
DR PIR; S16767; S16767.
DR PIR; S16771; S16771.
DR HSSP; O19673; 1HSB.
DR MIM; 142800; -
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003597; IG_C1.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; igf_1.
DR Pfam; PF00129; MHC_I; 1.
DR ProDom; PD000050; MHC_I; 1.
DR SMART; SM00407; IGC1; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal; Polymorphism.
FT SIGNAL 1 24
FT CHAIN 25 365
FT -----
FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT AW-34(A*10) ALPHA CHAIN.
FT -----
FT DOMAIN 25 114
FT DOMAIN 115 206
FT DOMAIN 207 298
FT DOMAIN 299 308
FT TRANSMEM 309 332
FT DOMAIN 333 365
FT CARBOHYD 110 110
FT DISULFID 125 188
FT DISULFID 227 283
FT VARIANT 3 3
FT VARIANT 90 90
FT VARIANT 121 121
FT VARIANT 129 129
FT VARIANT 138 138
FT VARIANT 180 180
FT VARIANT 312 312
FT SEQUENCE 365 AA; 41055 MW; 063BF63E5E6E01F6 CRC64;
SQ
Query Match 53.7%; Score 36; DB 1; Length 365;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 4 WXRNRKRV 12
Db 84 WDRNRKRV 92
RESULT 11
SP11-MYXVL
ID SP11-MYXVL STANDARD; PRT; 369 AA.

```

---

```

AC P12393;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serine proteinase inhibitor 1 (Serp1) (Serp-1).
GN SERP1 OR MO08.1L.
OS Myxoma virus (strain Lausanne).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Leporipoxvirus.
OX NCBI_TaxID=31530;
RN [1]
RX MEDLINE=91049428; PubMed=2173255;
RA Upton C., Macen J.L., Wisbart D.S., McPadden G.;
RT "Myxoma virus and malignant rabbit fibroma virus encode a serpin-like
RT protein important for virus virulence.";
RL Virology 179:618-631(1990).
RN [2]
RX MEDLINE=87064296; PubMed=3023828;
RA Upton C., McPadden G.;
RT "DNA sequence homology between the terminal inverted repeats of Shope
RT fibroma virus and an endogenous cellular plasmid species.";
RL Mol. Cell. Biol. 6:265-276(1986).
RN [3]
RX MEDLINE=87030884; PubMed=3021526;
RA Upton C., Carrell R.W., McPadden G.;
RT "A novel member of the serpin superfamily is encoded on a circular
RT plasmid-like DNA species isolated from rabbit cells.";
RL FEBS Lett. 207:115-120(1986).
RN [4]
RX MEDLINE=20032073; PubMed=10562494;
RA Cameron C., Hota-Mitchell S., Chen L., Barrett J., Cao J.X.,
RA Macaulay C., Willer D., Evans D., McPadden G.;
RT "The complete DNA sequence of Myxoma virus.";
RL Virology 264:298-318(1999).
CC -1- FUNCTION: IMPORTANT IN VIRULENCE.
CC -1- SIMILARITY: BELONGS TO THE SERPIN FAMILY. POXVIRUSES SUBFAMILY.
CC -1- CAUTION: WAS ORIGINALLY (REF.2 AND REF.3) THOUGHT TO ORIGINATE
CC FROM A PLASMID RABBIT DNA. THE ORIGINAL SAMPLE WAS CONTAMINATED
CC AND THE GENE IS DERIVED FROM MYXOMA VIRUS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M35233; AAA46629.1; -
DR EMBL; M12333; AAA81567.1; -
DR EMBL; AF170726; AAF15055.1; -
DR EMBL; AF170726; AAF14896.1; -
DR PIR; A24470; A24470.
DR PIR; B36418; B36418.
DR HSSP; P05121; 1B3K.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00079; serpin; 1.
DR SMART; SM00093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Serpin; Serine protease inhibitor; Virulence.
FT ACT_SITE 319 320 REACTIVE_BOND (BY SIMILARITY).
FT CONFLICT 1 16 MKYLVLCVLSACR -> MFWVRV (IN REF. 2
FT AND 3).
FT SEQUENCE 369 AA; 41556 MW; 8DB31CE131C218A0 CRC64;
SQ
Query Match 53.7%; Score 36; DB 1; Length 369;
Best Local Similarity 85.7%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```



```

Y 4 WXRNRK 10
b 235 WXRNRK 241

RESULT 12
YH CHLMU
D SYH CHLMU STANDARD; PRT; 428 AA.
C QPJJ9;
T 16-OCT-2001 (Rel. 40, Created)
T 16-OCT-2001 (Rel. 40, Last sequence update)
T 16-OCT-2001 (Rel. 40, Last annotation update)
E Histidyl-tRNA synthetase (EC 6.1.1.21) (Histidine--tRNA ligase)
E (Hiss).
N N HISS OR TC0830.
S Chlamydia muridarum.
C Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
X NCBI_TaxID=83560;
N N [1]
P SEQUENCE FROM N.A.
C STRAIN=MoPn / Ni99;
X MEDLINE=20150255; PubMed=10684935;
A Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
A White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
A Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
A Winn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
A Eisen J., Fraser C.M.;
T "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
T pneumoniae AR39."; 28:1397-1406(2000).
L Nucleic Acids Res.
C -1- CATALYTIC ACTIVITY: ATP + L-histidine + tRNA(His) = AMP +
C diposphate + L-histidyl-tRNA(His).
C -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
C -1- SUBCELLULAR LOCATION: Cytoplasmic.
C -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-tRNA SYNTHETASE FAMILY.
C This SWISS-PROT entry is copyright. It is produced through a collaboration
C between the Swiss Institute of Bioinformatics and the EMBL outstation -
C the European Bioinformatics Institute. There are no restrictions on its
C use by non-profit institutions as long as its content is in no way
C modified and this statement is not removed. Usage by and for commercial
C entities requires a license agreement (See http://www.isb-sib.ch/announce/
C or send an email to license@isb-sib.ch).
C EMBL; AE002349; AAP39630.1; -
R HSP; O32422; IQE0.
R TIGR; TC0830; -
R InterPro; IPR002106; AATRNA_ligaseII.
R InterPro; IPR004154; HGTP_anticonon.
R InterPro; IPR004516; Hiss.
R Pfam; PF00587; tRNA-synt_2b; 1.
R Pfam; PF03129; HGTP_anticonon; 1.
R TIGRFAMs; TIGR00442; hisS; 1.
R POSITE; PS50862; AA_tRNA_LIGASE_II; 1.
W Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
W Complete proteome.
Q SEQUENCE 428 AA; 48939 MW; 9CF859ED0E689DDF CRC64;

Query Match 53.7%; Score 36; DB 1; Length 428;
Best Local Similarity 44.4%; Pred. No. 21;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Y 1 CFQWXRNR 9
b 348 CFSWAKHLR 356

RESULT 13
P4 ROTGA
D VP4 ROTGA STANDARD; PRT; 749 AA.
C Q04916;
T 01-JUN-1994 (Rel. 29, Created)

US-09-743-107b-88.rsp
01-JUN-1994 (Rel. 29, Last sequence update)
01-JUN-1994 (Rel. 29, Last annotation update)
Outer capsid protein VP4 (Hemagglutinin) (Outer layer protein VP4).
S4.
OS Rotavirus (group B / strain ADRV) (Adult diarrhea rotavirus).
OC Viruses; dsRNA viruses; Reoviridae; Rotavirus.
OX NCBI_TaxID=12705;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93233240; PubMed=8396274;
RA Mackow E.R., Werner-Eckert R., Fay M.E., Tao H., Chen G.-M.;
RT "Identification and baculovirus expression of the VP4 protein of the
RT human group B rotavirus ADRV.";
RL J. Virol. 67:2730-2738(1993).
CC -1- SUBCELLULAR LOCATION: Outer capsid.
CC -1- SIMILARITY: BELONGS TO THE ROTAVIRUS VP4 PROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; M91434; AAA47338.1; -
DR InterPro; IPR000416; Cap_VP4.
DR Pfam; PF00426; VP4; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 29 29 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 53 53 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 109 109 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 133 133 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 407 407 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 527 527 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 568 568 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 681 681 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 698 698 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 749 AA; 84362 MW; D1223527DEAE0F21 CRC64;

Query Match 53.7%; Score 36; DB 1; Length 749;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CFQWXRNRKVR 12
Db 195 CFTWDMNCANVR 206

RESULT 14
RL28 THETN
ID RL28 THETN STANDARD; PRT; 62 AA.
AC Q8R9U1;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L28.
GN RPMB OR TTE1495.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ME4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
CC -1- SIMILARITY: BELONGS TO THE L28P FAMILY OF RIBOSOMAL PROTEINS.

```

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; A5013107; AM24713.1; -  
KW Ribosomal protein; Complete proteome.  
SQ SEQUENCE 62 AA; 7037 MW; CAADB605C81D495D CRC64;

Query Match 52.2%; Score 35; DB 1; Length 62;  
Best Local Similarity 60.0%; Pred. No. 4.6;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 QWXRNRKVR 12  
Db 27 RWKPNIRKVR 36

RESULT 15

VIA\_BMV VIA\_BMV STANDARD; PRT; 966 AA.  
AC Q00020;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE 1A protein [includes: Helicase; Methyltransferase].  
OS Broad bean mottle virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;  
OC Bromovirus.  
OX NCBI\_TaxID=12301;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bawden;  
RX MEDLINE=92074218; PubMed=1962437;  
RA Dzanott A.M., Bujarski J.J.;  
RT "The nucleotide sequence and genome organization of the RNA-1 segment  
RT in two bromoviruses: broad bean mottle virus and cowpea chlorotic  
RT mottle virus."  
RL Virology 185:553-562(1991).  
CC -!- FUNCTION: MAY BE INVOLVED IN THE REPLICATION OF THE VIRUS.  
CC CONTAINS AN HELICASE DOMAIN AND A METHYLTRANSFERASE DOMAIN. THE  
CC METHYLTRANSFERASE DOMAIN IS PROBABLY INVOLVED IN VIRAL RNA  
CC CAPPING.  
CC -!- SIMILARITY: TO 1A PROTEIN FROM CCMV, CMV, PSV AND TAV.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; M65138; AAA42740.1; -  
DR PIR; A41699; PIBVBE.  
DR InterPro; IPR002588; V\_methyltransf.  
DR InterPro; IPR000606; Viral\_helicase1.  
DR Pfam; PF01443; Viral\_helicase1.  
DR Pfam; PF01660; Vmethyltransf.1.  
KW Helicase; ATP-binding; Transferase; Methyltransferase.  
FT NP\_BIND 690 697 ATP (POTENTIAL).  
SQ SEQUENCE 966 AA; 109621 MW; DF592881D7231C8D CRC64;

Query Match 52.2%; Score 35; DB 1; Length 966;  
Best Local Similarity 46.7%; Pred. No. 72;  
Matches 7; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 1 CFQ----WXRNRKVR 11  
||:|:|:|

Db 347 CFKKNKDWTEENRSV 361  
Search completed: February 21, 2003, 07:51:36  
Job time : 6.2093 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:44:43 ; Search time 21.6744 Seconds  
(without alignments)  
114.078 Million cell updates/sec

Title: US-09-743-107B-88

Perfect score: 67

Sequence: 1 CFQWXRNRKV 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 571580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	89.6	711	4	Q8TCD2
2	56	83.6	38	4	Q9UCY5
3	45	67.2	33	6	Q9TR80
4	41	61.2	511	16	Q8Z462
5	40	59.7	148	10	Q9XHP1
6	40	59.7	273	2	Q31090
7	39	58.2	275	5	Q93780
8	39	58.2	341	11	Q8R2A4
9	39	58.2	415	5	Q9U4M9
10	39	58.2	866	10	Q9FHI9
11	38	56.7	108	9	Q8SC55
12	38	56.7	372	10	Q81553
13	38	56.7	531	10	Q9SFC0
14	37	55.2	91	15	Q77855
15	37	55.2	91	15	Q77856
16	37	55.2	179	11	Q9CZQ7

17	37	55.2	233	11	Q9D4X5	Q9d4x5 mus musculu
18	37	55.2	274	4	Q96M21	Q96m21 homo sapien
19	37	55.2	279	16	Q8XSE2	Q8xse2 ralestonia s
20	37	55.2	329	12	Q9QB73	Q9qb73 yaba monkey
21	37	55.2	368	11	Q9DAJ3	Q9daj3 mus musculu
22	37	55.2	368	11	Q9D9X0	Q9d9x0 mus musculu
23	37	55.2	368	11	Q9JIT1	Q9jit1 mus musculu
24	37	55.2	377	12	Q9LMQ5	Q9lmq5 lumpy skin
25	37	55.2	381	12	Q9DHK5	Q9dhk5 yaba-like d
26	37	55.2	489	16	P96223	P96223 mycobacteri
27	37	55.2	501	3	Q9UUD0	Q9uud0 schizosacch
28	37	55.2	570	10	Q8A487	Q8a487 zea mays (m
29	37	55.2	584	17	Q9HPA3	Q9hpa3 halobacteri
30	37	55.2	2348	5	Q9V346	Q9v346 drosophila
31	36	53.7	105	10	Q9AFD5	Q9afd5 oryza sativ
32	36	53.7	109	15	Q9YQC1	Q9yqc1 human immun
33	36	53.7	109	15	Q9YQC0	Q9yqc0 human immun
34	36	53.7	109	15	Q9YQB9	Q9yqb9 human immun
35	36	53.7	109	15	Q9YQB8	Q9yqb8 human immun
36	36	53.7	109	15	Q9YJ17	Q9yj17 human immun
37	36	53.7	109	15	Q9YJ32	Q9yj32 human immun
38	36	53.7	119	15	Q8Q454	Q8q454 human immun
39	36	53.7	124	10	Q9SYH0	Q9syh0 arabidopsis
40	36	53.7	298	16	Q8YF77	Q8yfh77 anabaena sp
41	36	53.7	306	4	Q8TAX2	Q8tax2 homo sapien
42	36	53.7	466	4	Q9NUS2	Q9nus2 homo sapien
43	36	53.7	514	5	Q9V118	Q9v118 drosophila
44	36	53.7	514	10	Q9SP27	Q9sp27 callistephu
45	36	53.7	543	5	Q9XZ30	Q9xz30 drosophila

#### ALIGNMENTS

#### RESULT 1

Q8TCD2 PRELIMINARY; PRT; 711 AA.  
AC Q8TCD2;  
DT 01-JUN-2002 (TRENBLrel. 21, Created)  
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
DE Lactotransferrin.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_taxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PROSTATE;  
RA Strauberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC022347; AAH2347.1; -  
SQ SEQUENCE 711 AA; 78327 MW; 1B9C7EE097C45FAF CRC64;

Query Match 89.6%; Score 60; DB 4; Length 711;  
Best Local Similarity 90.9%; Pred. No. 0.003; 1; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 1;

QY 1 CFQWXRNRKV 11  
|||  
Db 39 CFQWXRNRKV 49

#### RESULT 2

ID Q9UCY5 PRELIMINARY; PRT; 38 AA.  
AC Q9UCY5;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
DE Lactoferrin homolog (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]\_SEQUENCE FROM N.A.  
 RP MEDLINE=96081613; PubMed=8551695;  
 RX Sato I.;  
 RA "Characterization of the 84-kDa protein with ABH activity in human  
 RT seminal plasma";  
 RL Jpn. J. Legal Med. 49:281-293 (1995).  
 DR HSSP; P02788; 18KA.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 1.  
 SQ SEQUENCE 38 AA; 4459 MW; 0402F490B5EBDEB CRC64;

Query Match 83.6%; Score 56; DB 4; Length 38;  
 Best Local Similarity 90.9%; Pred. No. 0.0085;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 FQWXRNRKVR 12  
 ||| |||||  
 DB 21 FQWQRNRKVR 31

## RESULT 3

OYTR80  
 ID Q9TR80 PRELIMINARY; PRT; 33 AA.  
 AC Q9TR80;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Lactoferrin (fragment).  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]\_SEQUENCE  
 RP SEQUENCE  
 RX MEDLINE=95127729; PubMed=7827104;  
 RA Qian Z.Y., Jollies P., Migliore-Samur D., Fiat A.M.;  
 RL Blochim. Biophys. Acta 1243:125-32 (1995).  
 DR HSSP; O77698; 1CE2.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 1.  
 SQ SEQUENCE 33 AA; 3914 MW; D1904CAE15A73961 CRC64;

Query Match 67.2%; Score 45; DB 6; Length 33;  
 Best Local Similarity 54.5%; Pred. No. 0.098;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 CFQWXRNRKV 11  
 ||| |||||  
 DB 19 CYQWQKVRKL 29

## RESULT 4

OY2462  
 ID Q8Z462 PRELIMINARY; PRT; 511 AA.  
 AC Q8Z462;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical protein STY3070.  
 GN STY3070.

OS Salmonella typhi.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Salmonella.  
 OX NCBI\_TaxID=601;  
 RN [1]\_SEQUENCE FROM N.A.  
 RP STRAIN=CT18;  
 RX MEDLINE=21534947; PubMed=11677608;  
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,

RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,  
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,  
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,  
 RA Krogan A., Larsen T.S., Leather S., Moulis S., O'Garra P., Parry C.,  
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,  
 RA Whitehead S., Barrett B.G.;  
 RT "Complete genome sequence of a multiple drug resistant Salmonella  
 RL enterica serovar Typhi CT18";  
 RL Nature 413:848-852 (2001).  
 DR EMBL; AL627276; CAD06049.1; -  
 DR Hypothetical protein; Complete proteome  
 KW SEQUENCE 511 AA; 58126 MW; E2DD124510D178B CRC64;

Query Match 61.2%; Score 41; DB 16; Length 511;  
 Best Local Similarity 58.3%; Pred. No. 9.9;  
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 CFQWXRNRKVR 12  
 ||| |||||  
 DB 350 CFANDKVR 361

## RESULT 5

OYXHP1  
 ID Q9XHP1 PRELIMINARY; PRT; 148 AA.  
 AC Q9XHP1;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE 2S albumin.  
 OS Sesamum indicum (Oriental sesame) (Gingelly).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Lamiales; Pedaliaceae; Sesamum.  
 OX NCBI\_TaxID=4182;  
 RN [1]\_SEQUENCE FROM N.A.  
 RP SEQUENCE  
 RC STRAIN=TAINAN 1;  
 RX MEDLINE=20074970; PubMed=10606554;  
 RA Tai S.S.K., Wu L.S.H., Chen E.C.F., Tzen J.T.C.;  
 RL "Molecular cloning of 11S globulin and 2S albumin, the two major seed  
 RT storage proteins in sesame";  
 RL J. Agric. Food Chem. 47:4932-4938 (1999).  
 DR EMBL; AF091841; AAD42943.1; -  
 DR InterPro; IPR003612; AAI.  
 DR InterPro; IPR000817; Napin.  
 DR InterPro; IPR001768; Try/amyL\_inhbr.  
 DR Pfam; PF00234; tryp\_alpha\_aml; 1.  
 DR PRINTS; PR00496; NAFIN.  
 DR SMART; SM00499; AAI; 1.  
 SQ SEQUENCE 148 AA; 17524 MW; BA46B033BA13E3DE CRC64;

Query Match 59.7%; Score 40; DB 10; Length 148;  
 Best Local Similarity 66.7%; Pred. No. 4.3;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CFQWXRNRMR 9  
 ||| |||||  
 DB 54 CQWQWESMR 62

## RESULT 6

OY1090  
 ID O31090 PRELIMINARY; PRT; 273 AA.  
 AC O31090;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
 DE Hypothetical 31.0 kDa protein.  
 OS Rhizobium leguminosarum (biovar viciae).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.

```

X NCBI_TaxID=387;
N SEQUENCE FROM N.A.
P STRAIN=VF39;
C Ivashina T.V., Sadykov M.R., Kanapin A.A., Ksenzenko V.N.;
L Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
N [2]
P SEQUENCE FROM N.A.
C STRAIN=VF39;
X MEDLINE=99113394; PubMed=9914965;
A Sadykov M.R., Ivashina T.V., Kanapin A.A., Shlyapnikov M.G.,
A Ksenzenko V.N.;
T "Structural and functional organization of the exopolysaccharide
T biosynthesis genes in Rhizobium leguminosarum bv. viciae VF39."
L Mol. Biol. (Mosk) 32:797-804 (1998).
R EMBL; AF028810; AAB8891.1; -.
W Hypothetical protein.
Q SEQUENCE 273 AA; 30986 MW; F195D2FEF7AD44D3 CRC64;

Query Match 59.7%; Score 40; DB 2; Length 273;
Best Local Similarity 60.0%; Pred. No. 8.1;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Y 3 QWXRNMKVR 12
:|:|:|:|:|
b 245 RWLRNLRKLR 254

RESULT 7
93780
D Q33780 PRELIMINARY; PRT; 275 AA.
C Q33780;
T 01-FEB-1997 (TrEMBLrel. 02, Created)
T 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
T 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
E F53H4.4 protein.
N F53H4.4.
S Caenorhabditis elegans.
C Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
C Rhabditidae; Peloderinae; Caenorhabditis.
X NCBI_TaxID=6239;
N [1]
P SEQUENCE FROM N.A.
A Dobson R.;
L Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
N [2]
P SEQUENCE FROM N.A.
X MEDLINE=99069613; PubMed=9851916;
A none;
T "Genome sequence of the nematode C.elegans: A platform for
T investigating biology."
L Science 282:2012-2018 (1998).
R EMBL; Z81089; CAB03137.1; -.
Q SEQUENCE 275 AA; 31716 MW; 9D69524FFC704DB2 CRC64;

Query Match 58.2%; Score 39; DB 5; Length 275;
Best Local Similarity 63.6%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Y 2 FQWXRNMKVR 12
:|:|:|:|:|
b 262 FQWXRNMKVR 272

RESULT 8
8R2A4
D Q8R2A4 PRELIMINARY; PRT; 341 AA.
C Q8R2A4;
T 01-JUN-2002 (TrEMBLrel. 21, Created)
T 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
T 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
E Vomeroneasal receptor V1RE6.
N V1RE6.

```

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129X1/SVJ;
RX MEDLINE=21676859; PubMed=11802169;
RA Rodriguez I., Del Punta K., Rothman A., Ishii T., Mombaerts P.;
RT "Multiple new and isolated families within the mouse superfamily of
RT Vir vomeronasal receptors."
RL Nat. Neurosci. 5:134-140 (2002).
DR EMBL; AY065506; AAL47911.1; -.
XW Receptor.
SQ SEQUENCE 341 AA; 39632 MW; B3D8259F7AD11130 CRC64;

Query Match 58.2%; Score 39; DB 11; Length 341;
Best Local Similarity 55.6%; Pred. No. 16;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWXRNMVR 9
:|:|:|:|:|
Db 331 CFQWXRNMVR 339

RESULT 9
Q9U4M9
ID Q9U4M9 PRELIMINARY; PRT; 415 AA.
AC Q9U4M9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 7138.2.
GN 7138.2.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RA Ravel C., Dubessay P., Blaineau C., Lignon M.-F., Bastien P.,
RA Dedet J.-P., Pages M.;
RT "Leishmania major chromosome 5 complete sequence."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF163772; AAF14642.1; -.
DR InterPro: IPR003409; MORN.
DR Pfam: PF02493; MORN; 4.
SQ SEQUENCE 415 AA; 46701 MW; 3E3AD710BF23691E CRC64;

Query Match 58.2%; Score 39; DB 5; Length 415;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWXRNMVR 9
:|:|:|:|:|
Db 272 CGEWFNRMR 280

RESULT 10
Q9FH19
ID Q9FH19 PRELIMINARY; PRT; 866 AA.
AC Q9FH19;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Genomic DNA, chromosome 5, p1 clone:MFC19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN=COLUMBIA;  
RX MEDLINE=99397451; PubMed=10470850;  
RA Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,  
RA Miyajima N., Tabata S.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 5. IX.  
RT Sequence features of the regions of 1,011,550 bp covered by seventeen  
RT P1 and TAC clones.";  
RL DNA Res. 6:183-195(1999).  
DR EMBL; AB018113; BAB09175.1; -;  
SQ SEQUENCE 866 AA; 100187 MW; 164FAB9B4976D27D CRC64;  
  
Query Match 58.2%; Score 39; DB 10; Length 866;  
Best Local Similarity 50.0%; Pred. No. 42;  
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
  
QY 1 CFQWXRNRKVR 12  
Db 389 CFNWLKFRKLR 400  
|||:|  
|||:|  
  
RESULT 11  
Q8SC55 PRELIMINARY; PRT; 108 AA.  
AC Q8SC55;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Hypothetical 12.3 kDa protein.  
OS Stx2 converting bacteriophage I.  
OC Viruses.  
OX NCBI\_TaxID=180816;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=STX2 PHAGE-I;  
RA Sato T., Shimizu T., Watarai M., Kobayashi M., Kano S., Hamabata T.,  
RA Yamasaki S., Takeda Y.;  
RT "Genomic sequence of Shiga toxin 2-converting phage isolated from  
RT Escherichia coli O157:H7 Okayama strain and comparison with other  
RT Shiga toxin 2-converting phages.";  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF004402; BAB87947.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 108 AA; 12279 MW; 2FF2E76690C78BA7 CRC64;  
  
Query Match 56.7%; Score 38; DB 9; Length 108;  
Best Local Similarity 60.0%; Pred. No. 7.5;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 CFQWXRNRK 10  
Db 31 CFQWISNRK 40  
|||:|  
|||:|  
  
RESULT 12  
Q81653 PRELIMINARY; PRT; 372 AA.  
AC Q81653;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Senescence-associated protein 3 (Fragment).  
GN SA3.  
OS Hemerocallis hybrid cultivar.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;  
OC Hemerocallidaceae; Hemerocallis.  
OX NCBI\_TaxID=80862;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. STELLA D'ORO; TISSUE=SENESCING PETALS;  
RX MEDLINE=99339248; PubMed=10412903;  
RA Panavas T., Pikula A., Reid P.D., Rubinstein B., Walker E.L.;  
RT "Identification of senescence-associated genes from daylily petals.";

RL Plant Mol. Biol. 40:237-248(1999).  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
DR EMBL; AF082028; AAC34853.1; -;  
DR InterPro; IPR001128; Cytochrome\_P450.  
DR Pfam; PF00067; P450; 1.  
DR PRINTS; PRO0385; P450.  
DR PROSITE; PS00086; CYTOCHROME\_P450; UNKNOWN\_1.  
KW Heme; Monooxygenase; Oxidoreductase.  
FT NON TER 1  
SQ SEQUENCE 372 AA; 41909 MW; 55FB3EAF96959595E CRC64;  
  
Query Match 56.7%; Score 38; DB 10; Length 372;  
Best Local Similarity 71.4%; Pred. No. 27;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 CFQWXRN 7  
Db 328 CFWDNRN 334  
|||:|  
|||:|  
  
RESULT 13  
Q9SIC0 PRELIMINARY; PRT; 531 AA.  
AC Q9SIC0;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE En/Spm-like transposon protein.  
GN AT2G05650.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosidis II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RX MEDLINE=20083487; PubMed=10617197;  
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,  
RA Buel C.R., Ketchum K.A., Lee J.J., Renning C.M., Koo H., Moffat K.S.,  
RA Cronin L.A., Shen M., VanAken S.B., Umayam L., Tallon L.J., Gill J.E.,  
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,  
RA Copenhaver G.P., Preuss D., Nietman W.C., White O., Eisen J.A.,  
RA Salzberg S.L., Fraser C.M., Venter J.C.;  
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
RT thaliana.";  
RL Nature 402:761-768(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA Lin X.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC007170; AAD25641.1; -;  
DR InterPro; IPR004264; Transposase\_23.  
DR Pfam; PF03017; Transposase\_23; 1.  
SQ SEQUENCE 531 AA; 60512 MW; 57B3AC60C976AAB9 CRC64;  
  
Query Match 56.7%; Score 38; DB 10; Length 531;  
Best Local Similarity 75.0%; Pred. No. 39;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 3 QWXRNRK 10  
Db 501 QWFRNRK 508  
|||:|  
|||:|  
  
RESULT 14  
Q77855 PRELIMINARY; PRT; 91 AA.  
AC Q77855;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

```

JT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
JE Envelope protein (Fragment).
JN ENV.
JS Human immunodeficiency virus type 1.
JC Viruses; Retroid viruses; Retroviridae; Lentivirus.
JX NCBI_TaxID=11676;
JN [1]
JP SEQUENCE FROM N.A.
JX MEDLINE=95191002; PubMed=7884875;
JA Mulder-Kampinga G.A., Simonon A., Kuiken C.L., Dekker J.,
JA Scherpbier H.J., de Perre P., Boer K., Goudsmit J.;
JT "Similarity in env and gag genes between genomic RNAs of human
JT immunodeficiency virus type 1 (HIV-1) from mother and infant is
JT unrelated to time of HIV-1 RNA positivity in the child.";
JL J. Virol. 69:2285-2296(1995).
JR EMBL; Z47867; CAA87881.1; -.
JR InterPro; IPR000777; GP120.
JR Pfam; PF00516; GP120; 1.
JW AIDS; Coat protein; Glycoprotein.
JT NON_TER 1
JT NON_TER 91
JQ SEQUENCE 91 AA; 10530 MW; 8B10C62011F305D6 CRC64;

Query Match 55.2%; Score 37; DB 15; Length 91;
Best Local Similarity 60.0%; Pred. No. 9.9;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

NY 3 QWXRNMKVR 12
  |||:|
  67 QWRTLQKVR 76

RESULT 15
Y77856 PRELIMINARY; PRT; 91 AA.
YC Q77856;
YT 01-NOV-1996 (TrEMBLrel. 01, Created)
YT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
YT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
JE Envelope protein (Fragment).
JN ENV.
JS Human immunodeficiency virus type 1.
JC Viruses; Retroid viruses; Retroviridae; Lentivirus.
JX NCBI_TaxID=11676;
JN [1]
JP SEQUENCE FROM N.A.
JX MEDLINE=95191002; PubMed=7884875;
JA Mulder-Kampinga G.A., Simonon A., Kuiken C.L., Dekker J.,
JA Scherpbier H.J., de Perre P., Boer K., Goudsmit J.;
JT "Similarity in env and gag genes between genomic RNAs of human
JT immunodeficiency virus type 1 (HIV-1) from mother and infant is
JT unrelated to time of HIV-1 RNA positivity in the child.";
JL J. Virol. 69:2285-2296(1995).
JR EMBL; Z47868; CAA87882.1; -.
JR InterPro; IPR000777; GP120.
JR Pfam; PF00516; GP120; 1.
JW AIDS; Coat protein; Glycoprotein.
JT NON_TER 1
JT NON_TER 91
JQ SEQUENCE 91 AA; 10519 MW; EB20C63A22DA1288 CRC64;

Query Match 55.2%; Score 37; DB 15; Length 91;
Best Local Similarity 60.0%; Pred. No. 9.9;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

NY 3 QWXRNMKVR 12
  |||:|
  67 QWRTLQKVR 76
```

Search completed: February 21, 2003, 08:00:43  
Ob time : 22.6744 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

DM protein - protein search, using sw model

run on: February 21, 2003, 07:37:21 ; Search time 28.093 Seconds  
(without alignments)  
56.918 Million cell updates/sec

Title: US-09-743-107B-89

Perfect score: 66

Sequence: 1 CFQWQRRXKVR 12

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq 101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	98.5	12	21 AAY78074	Human lactoferrin
2	64	97.0	12	21 AAY78038	Human lactoferrin
3	64	97.0	12	21 AAY78046	Human lactoferrin
4	64	97.0	12	21 AAY78047	Human lactoferrin
5	64	97.0	12	21 AAY78086	Human lactoferrin
6	64	97.0	12	21 AAY78089	Human lactoferrin
7	64	97.0	12	21 AAY78090	Human lactoferrin
8	64	97.0	13	21 AAY78037	Human lactoferrin
9	64	97.0	13	21 AAY78048	Human lactoferrin
10	64	97.0	13	21 AAY78049	Human lactoferrin

11	64	97.0	14	21 AAY78036	Human lactoferrin
12	64	97.0	14	21 AAY78050	Human lactoferrin
13	64	97.0	14	21 AAY78051	Human lactoferrin
14	64	97.0	15	17 AAR98554	Peptide for anti-u
15	64	97.0	15	21 AAY78035	Human lactoferrin
16	64	97.0	15	21 AAY78062	Human lactoferrin
17	64	97.0	15	21 AAY78063	Human lactoferrin
18	64	97.0	16	21 AAY78031	Human lactoferrin
19	64	97.0	16	21 AAY78064	Human lactoferrin
20	64	97.0	16	21 AAY78065	Human lactoferrin
21	64	97.0	17	21 AAY78034	Human lactoferrin
22	64	97.0	17	21 AAY78066	Human lactoferrin
23	64	97.0	17	21 AAY78067	Human lactoferrin
24	64	97.0	18	15 AAR69352	Human lactoferrin
25	64	97.0	18	17 AAW13397	Advanced glycosyla
26	64	97.0	18	21 AAY78033	Human lactoferrin
27	64	97.0	19	21 AAY68867	Amino acid sequenc
28	64	97.0	19	21 AAY78032	Human lactoferrin
29	64	97.0	20	13 AAR21810	Anti microbial pep
30	64	97.0	20	14 AAR44841	Lactoferrin-relate
31	64	97.0	20	15 AAR48530	Lactoferrin derive
32	64	97.0	20	15 AAR48531	Lactoferrin derive
33	64	97.0	20	15 AAR57461	Lactoferrin derive
34	64	97.0	20	15 AAR57462	Lactoferrin derive
35	64	97.0	20	16 AAR84698	Bovine lactoferrin
36	64	97.0	20	16 AAR84699	Bovine lactoferrin
37	64	97.0	20	16 AAR80263	Anti-parasitic lac
38	64	97.0	20	16 AAR80264	Anti-parasitic lac
39	64	97.0	20	17 AAR98553	Peptide for anti-u
40	64	97.0	20	17 AAR91852	Lactoferrin-derive
41	64	97.0	20	17 AAW03045	Lactoferrin-derive
42	64	97.0	20	17 AAR90607	Lactoferrin-derive
43	64	97.0	20	17 AAR87621	Lactoferrin-derive
44	64	97.0	20	17 AAR87622	Lactoferrin-derive
45	64	97.0	20	18 AAW26150	Lactoferrin deriva

## ALIGNMENTS

### RESULT 1

AAV78074

ID AAY78074 standard; Peptide; 12 AA.

AC AAY78074;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:74.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
XX urinary tract infection; colitis; Candida infection; fungicidal;  
XX bactericidal; preservative.

OS Homo sapiens.

XX Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SB01230.

XX 06-JUL-1998; 98SE-0002441.

XX 17-JUL-1998; 98SE-0002562.

XX 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.



XX	New peptides used for treatment and prevention of infections,
PT	inflammations and tumors and for use in infant formula food -
XX	Claim 22; Page 35; 102pp; English.
XX	AAAY78001 to AAAY78100 represent peptides having sequences based on human
CC	lactoferrin. The peptides are taken up in the intestine through
CC	binding to specific lactoferrin receptors and are then transported
CC	through the circulation. A medicinal product of the peptide or fragment
CC	can be used for treating and/or prevention of infections (such as
CC	urinary tract infections, colitis, and Candida infection on a mucosal
CC	membrane), inflammations and/or tumours. The peptides can also be used
CC	in food stuffs such as infant formula food. The peptides are also
CC	fungicidal and bactericidal and may also be used as preservatives.
CC	Even though native human lactoferrin have been shown to have desired
CC	anti-inflammatory anti-infectious and anti-tumoural properties they
CC	cannot be used clinically on a broad basis because of high production
CC	costs. Therefore, provision of peptides based on lactoferrin would
CC	enable them to be used for the same purposes as lactoferrin at lower
CC	cost.
XX	
XX	Sequence 12 AA;
SQ	
	Query Match 98.5%; Score 65; DB 21; Length 12;
	Best Local Similarity 91.7%; Pred. No. 5.3e-05;
	Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	1 CFOWQRXMRKVR 12
Db	1 CFOWQRXMRKVR 12
RESULT 2	
AAAY78038	
ID	AAAY78038 standard; Peptide; 12 AA.
XX	
AC	AAAY78038;
DT	25-APR-2000 (first entry)
DE	Human lactoferrin derived peptide SEQ ID NO:39.
XX	
KW	Human; lactoferrin; modification; infection; inflammation; tumour;
KW	food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;
KW	urinary tract infection; colitis; Candida infection; fungicidal;
KW	bactericidal; preservative.
OS	Homo sapiens.
OS	Synthetic.
FN	WO200001730-A1.
PD	13-JAN-2000.
XX	
PF	06-JUL-1999; 99WO-S001230.
XX	
PR	06-JUL-1998; 98SE-0002441.
PR	17-JUL-1998; 98SE-0002562.
PR	29-DEC-1998; 98SE-0004614.
XX	
PA	(ASCI-) A+ SCI INVEST AB.
XX	
PI	Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;
XX	
DR	WPI; 2000-147388/13.
XX	
PT	New peptides used for treatment and prevention of infections,
PT	inflammations and tumors and for use in infant formula food -
XX	Claim 12; Page 70; 102pp; English.
XX	AAAY78001 to AAAY78100 represent peptides having sequences based on human

lactoferrin. The peptides are taken up in the intestine through binding to specific lactoferrin receptors and are then transported through the circulation. A medicinal product of the peptide or fragment can be used for treating and/or prevention of infections (such as urinary tract infections, colitis, and Candida infection on a mucosal membrane), inflammations and/or tumours. The peptides can also be used in food stuffs such as infant formula food. The peptides are also fungicidal and bactericidal and may also be used as preservatives. Even though native human lactoferrin have been shown to have desired anti-inflammatory anti-infectious and anti-tumoural properties they cannot be used clinically on a broad basis because of high production costs. Therefore, provision of peptides based on lactoferrin would enable them to be used for the same purposes as lactoferrin at lower cost.

XX  
SQ Sequence 12 AA;

Query Match 97.0%; Score 64; DB 21; Length 12;  
Best Local Similarity 91.7%; Pred. No. 7.9e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRNMRKVR 12  
||| |||  
Db 1 CFQWQRNMRKVR 12

RESULT 3  
AAY78046  
ID AAY78046 standard; Peptide; 12 AA.  
XX  
AC AAY78046;  
DT  
DE 25-APR-2000 (first entry)  
DE Human lactoferrin derived peptide SEQ ID NO:46.  
XX  
KW Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
OS Homo sapiens.  
OS Synthetic.  
XX WO200001730-A1.  
FN  
PD 13-JAN-2000.  
XX  
PF 06-JUL-1999; 99WO-SB01230.  
PR 06-JUL-1998; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
XX  
PA (ASCI-) A+ SCI INVEST AB.  
XX  
XX Hanson LA, Mattesby-Baltzer I, Baltzer L, Dolphin GT;  
PI  
DR WPI; 2000-147388/13.  
XX  
PT New peptides used for treatment and prevention of infections, -  
FT Inflammations and tumors and for use in infant formula food  
XX  
XX Claim 15; Page 35; 102pp; English.  
XX  
XX AAY78001 to AAY78100 represent peptides having sequences based on human  
XX lactoferrin. The peptides are taken up in the intestine through  
XX binding to specific lactoferrin receptors and are then transported  
XX through the circulation. A medicinal product of the peptide or fragment  
XX can be used for treating and/or prevention of infections (such as  
XX urinary tract infections, colitis, and Candida infection on a mucosal  
XX membrane), inflammations and/or tumours. The peptides can also be used  
XX in food stuffs such as infant formula food. The peptides are also  
XX

XX fungicidal and bactericidal and may also be used as preservatives.  
 XX Even though native human lactoferrin have been shown to have desired  
 XX anti-inflammatory anti-infectious and anti-tumoural properties they  
 XX cannot be used clinically on a broad basis because of high production  
 XX costs. Therefore, provision of peptides based on lactoferrin would  
 XX enable them to be used for the same purposes as lactoferrin at lower  
 XX cost.

XX Sequence 12 AA;

Query Match 97.0%; Score 64; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 7.9e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

YY 1 CFQWQRXMRKVR 12  
 ||||| |||||  
 YY 1 CFQWQRXMRKVR 12

RESULT 4  
 AAY78047  
 ID AAY78047 standard; Peptide; 12 AA.

XX AAY78047;  
 XX  
 XX 25-APR-2000 (first entry)  
 XX Human lactoferrin derived peptide SEQ ID NO:47.  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX urinary tract infection; colitis; Candida infection; fungicidal;  
 XX bactericidal; preservative.

XX Homo sapiens.  
 XX Synthetic.  
 XX WO200001730-A1.  
 XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.  
 XX 06-JUL-1998; 98SE-0002441.  
 XX 17-JUL-1998; 98SE-0002562.  
 XX 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 XX inflammations and tumors and for use in infant formula food -  
 XX Claim 18; Page 73; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 XX lactoferrin. The peptides are taken up in the intestine through  
 XX binding to specific lactoferrin receptors and are then transported  
 XX through the circulation. A medicinal product of the peptide or fragment  
 XX can be used for treating and/or prevention of infections (such as  
 XX urinary tract infections, colitis, and Candida infection on a mucosal  
 XX membrane), inflammations and/or tumours. The peptides can also be used  
 XX in food stuffs such as infant formula food. The peptides are also  
 XX fungicidal and bactericidal and may also be used as preservatives.  
 XX Even though native human lactoferrin have been shown to have desired  
 XX anti-inflammatory anti-infectious and anti-tumoural properties they  
 XX cannot be used clinically on a broad basis because of high production  
 XX costs. Therefore, provision of peptides based on lactoferrin would  
 XX enable them to be used for the same purposes as lactoferrin at lower  
 XX cost.

XX SQ Sequence 12 AA;

Query Match 97.0%; Score 64; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 7.9e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CFQWQRXMRKVR 12  
 ||||| |||||  
 Db 1 CFQWQRXMRKVR 12

RESULT 5  
 AAY78086  
 ID AAY78086 standard; Peptide; 12 AA.

XX AAY78086;  
 XX  
 XX 25-APR-2000 (first entry)  
 XX Human lactoferrin derived peptide SEQ ID NO:86.  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX urinary tract infection; colitis; Candida infection; fungicidal;  
 XX bactericidal; preservative.

XX Homo sapiens.  
 XX Synthetic.  
 XX WO200001730-A1.  
 XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.  
 XX 06-JUL-1998; 98SE-0002441.  
 XX 17-JUL-1998; 98SE-0002562.  
 XX 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 XX inflammations and tumors and for use in infant formula food -  
 XX Claim 22; Page 36; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 XX lactoferrin. The peptides are taken up in the intestine through  
 XX binding to specific lactoferrin receptors and are then transported  
 XX through the circulation. A medicinal product of the peptide or fragment  
 XX can be used for treating and/or prevention of infections (such as  
 XX urinary tract infections, colitis, and Candida infection on a mucosal  
 XX membrane), inflammations and/or tumours. The peptides can also be used  
 XX in food stuffs such as infant formula food. The peptides are also  
 XX fungicidal and bactericidal and may also be used as preservatives.  
 XX Even though native human lactoferrin have been shown to have desired  
 XX anti-inflammatory anti-infectious and anti-tumoural properties they  
 XX cannot be used clinically on a broad basis because of high production  
 XX costs. Therefore, provision of peptides based on lactoferrin would  
 XX enable them to be used for the same purposes as lactoferrin at lower  
 XX cost.

XX Sequence 12 AA;

Query Match 97.0%; Score 64; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 7.9e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQXMRKVR 12  
 DB 1 CFQWQXMRKVR 12

## RESULT 6

AAV78089  
 ID AAY78089 standard; Peptide; 12 AA.

XX AAY78089;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:89.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX urinary tract infection; colitis; Candida infection; fungicidal;  
 XX bactericidal; preservative.

XX Homo sapiens.

OS Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

XX 17-JUL-1998; 98SE-0002562.

XX 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 XX inflammations and tumors and for use in infant formula food -

XX Claim 22; Page 37; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 12 AA;

Query Match 97.0%; Score 64; DB 21; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 7.9e-05;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQXMRKVR 12

DB 1 CFQWQXMRKVR 12

## RESULT 7

AAV78090

ID AAY78090 standard; Peptide; 12 AA.

XX AAY78090;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:90.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX urinary tract infection; colitis; Candida infection; fungicidal;  
 XX bactericidal; preservative.

XX Homo sapiens.

OS Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

XX 17-JUL-1998; 98SE-0002562.

XX 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 XX inflammations and tumors and for use in infant formula food -

XX Claim 22; Page 37; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 12 AA;

Query Match 97.0%; Score 64; DB 21; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 7.9e-05;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQXMRKVR 12

DB 1 CFQWQXMRKVR 12

## RESULT 8

AAV78037

ID AAY78037 standard; Peptide; 13 AA.

XX AAY78037;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:37.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 DS Synthetic.  
 XX WO200001730-A1.  
 PN 13-JAN-2000.  
 XX 06-JUL-1999; 99WO-SE01230.  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 WI WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 TT inflammations and tumors and for use in infant formula food -  
 XX Claim 12; Page 70; 102pp; English.  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX SQ Sequence 13 AA;  
 Query Match 97.0%; Score 64; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 8.6e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Y 1 CFQWQXMRKVR 12  
 |||||  
 b 2 CFQWQXMRKVR 13  
 RESULT 9  
 AAY78048  
 ID AAY78048 standard; Peptide; 13 AA.  
 XX AAY78048;  
 XX 25-APR-2000 (first entry)  
 XX Human lactoferrin derived peptide SEQ ID NO:48.  
 XX Homo sapiens.  
 XX Synthetic.  
 XX WO200001730-A1.  
 PN 13-JAN-2000.  
 XX 06-JUL-1999; 99WO-SE01230.  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 WI WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 TT inflammations and tumors and for use in infant formula food -  
 XX Claim 12; Page 70; 102pp; English.  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

OS Synthetic.  
 XX WO200001730-A1.  
 PN 13-JAN-2000.  
 XX 06-JUL-1999; 99WO-SE01230.  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 WI WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 TT inflammations and tumors and for use in infant formula food -  
 XX Claim 15; Page 74; 102pp; English.  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX SQ Sequence 13 AA;  
 Query Match 97.0%; Score 64; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 8.6e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Y 1 CFQWQXMRKVR 12  
 |||||  
 b 2 CFQWQXMRKVR 13  
 RESULT 10  
 AAY78049  
 ID AAY78049 standard; Peptide; 13 AA.  
 XX AAY78049;  
 XX 25-APR-2000 (first entry)  
 XX Human lactoferrin derived peptide SEQ ID NO:49.  
 XX Homo sapiens.  
 XX Synthetic.  
 XX WO200001730-A1.  
 PN 13-JAN-2000.  
 XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 XX Hanson LA, Matteby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 DR New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 18; Page 74; 102pp; English.  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX Sequence 13 AA;  
 SQ Query Match 97.0%; Score 64; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 8.6e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 CFQWQRMKVR 12  
 DB 2 CFQWQRMKVR 13  
 RESULT 11  
 AAY78036  
 ID AAY78036 standard; Peptide; 14 AA.  
 XX AAY78036;  
 XX 25-APR-2000 (first entry)  
 XX Human lactoferrin derived peptide SEQ ID NO:36.  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.  
 OS WO200001730-A1.  
 XX 13-JAN-2000.  
 XX 06-JUL-1998; 99WO-SE01230.  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 XX Hanson LA, Matteby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 DR New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 18; Page 74; 102pp; English.

PI Hanson LA, Matteby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 12; Page 69; 102pp; English.  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX Sequence 14 AA;  
 SQ Query Match 97.0%; Score 64; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 9.2e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 CFQWQRMKVR 12  
 DB 3 CFQWQRMKVR 14  
 RESULT 12  
 AAY78050  
 ID AAY78050 standard; Peptide; 14 AA.  
 XX AAY78050;  
 XX 25-APR-2000 (first entry)  
 XX Human lactoferrin derived peptide SEQ ID NO:50.  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.  
 OS WO200001730-A1.  
 XX 13-JAN-2000.  
 XX 06-JUL-1998; 99WO-SE01230.  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 XX Hanson LA, Matteby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 DR New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 12; Page 69; 102pp; English.

2S Claim 15; Page 75; 102pp; English.

3C AAY78001 to AAY78100 represent peptides having sequences based on human  
 3C lactoferrin. The peptides are taken up in the intestine through  
 3C binding to specific lactoferrin receptors and are then transported  
 3C through the circulation. A medicinal product of the peptide or fragment  
 3C can be used for treating and/or prevention of infections (such as  
 3C urinary tract infections, colitis, and Candida infection on a mucosal  
 3C membrane), inflammations and/or tumours. The peptides can also be used  
 3C in food stuffs such as infant formula food. The peptides are also  
 3C fungicidal and bactericidal and may also be used as preservatives.  
 3C Even though native human lactoferrin have been shown to have desired  
 3C anti-inflammatory anti-infectious and anti-tumoural properties they  
 3C cannot be used clinically on a broad basis because of high production  
 3C costs. Therefore, provision of peptides based on lactoferrin would  
 3C enable them to be used for the same purposes as lactoferrin at lower  
 3C cost.

3Q Sequence 14 AA;

Query Match 97.0%; Score 64; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 9.2e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

3Y 1 CFQWQRMKVR 12  
 ||||| |||||  
 3 CFQWQRMKVR 14

3b 3 CFQWQRMKVR 14

RESULT 13

3AY78051

3D AAY78051 standard; Peptide; 14 AA.

3E AAY78051;

3F 25-APR-2000 (first entry)

3G Human lactoferrin derived peptide SEQ ID NO:51.

3H Human; lactoferrin; modification; infection; inflammation; tumour;  
 3I food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 3J urinary tract infection; colitis; Candida infection; fungicidal;  
 3K bactericidal; preservative.

3L Homo sapiens.

3M Synthetic.

3N WO200001730-A1.

3O 13-JAN-2000.

3P 06-JUL-1999; 99WO-SE01230.2

3Q 06-JUL-1998; 98SE-0002441.

3R 17-JUL-1998; 98SE-0002562.

3S 29-DEC-1998; 98SE-0004614.

3T (ASCI-) A+ SCI INVEST AB.

3U Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 3V WPI; 2000-147388/13.

3W New peptides used for treatment and prevention of infections,  
 3X inflammations and tumors and for use in infant formula food -  
 3Y Claim 18; Page 75; 102pp; English.

3Z AAY78001 to AAY78100 represent peptides having sequences based on human  
 3C lactoferrin. The peptides are taken up in the intestine through  
 3C binding to specific lactoferrin receptors and are then transported  
 3C through the circulation. A medicinal product of the peptide or fragment  
 3C can be used for treating and/or prevention of infections (such as

3C urinary tract infections, colitis, and Candida infection on a mucosal  
 3C membrane), inflammations and/or tumours. The peptides can also be used  
 3C in food stuffs such as infant formula food. The peptides are also  
 3C fungicidal and bactericidal and may also be used as preservatives.  
 3C Even though native human lactoferrin have been shown to have desired  
 3C anti-inflammatory anti-infectious and anti-tumoural properties they  
 3C cannot be used clinically on a broad basis because of high production  
 3C costs. Therefore, provision of peptides based on lactoferrin would  
 3C enable them to be used for the same purposes as lactoferrin at lower  
 3C cost.

3Q Sequence 14 AA;

Query Match 97.0%; Score 64; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 9.2e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

3Y 1 CFQWQRMKVR 12  
 ||||| |||||  
 3 CFQWQRMKVR 14

3b 3 CFQWQRMKVR 14

RESULT 14

3AR98554

3D AAR98554 standard; Peptide; 15 AA.

3E AAR98554;

3F 12-NOV-1996 (first entry)

3G Peptide for anti-ulcer agent.

3H anti-ulcer agent; low toxicity; stable; heat-resistant.

3I Synthetic.

3J JP08143468-A.

3K 04-JUN-1996.

3L 17-NOV-1994; 94JP-0283869.

3M 17-NOV-1994; 94JP-0283869.

3N (MORG) MORINAGA MILK IND CO LTD.

3O WPI; 1996-318857/32.

3P Anti-ulcer agent contg. peptide - has low toxicity, is  
 3Q heat-resistant and water-soluble

3R Claim 1; Page 11; 11pp; Japanese.

3S AAR98531-54 are peptides used in an anti-ulcer agent. The agent is low  
 3T in toxicity, is heat-resistant and stable in aqueous soln.. It can be  
 3U administered orally and be produced in large amounts.

3V Sequence 15 AA;

Query Match 97.0%; Score 64; DB 17; Length 15;  
 Best Local Similarity 91.7%; Pred. No. 9.8e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

3Y 1 CFQWQRMKVR 12  
 ||||| |||||  
 2 CFQWQRMKVR 13

3b 2 CFQWQRMKVR 13

RESULT 15

3AY78035

3D AAY78035 standard; Peptide; 15 AA.

3E AAY78035;

XX	25-APR-2000	(first entry)	
DT	Human lactoferrin derived peptide	SEQ ID NO:35.	
XX	Human; lactoferrin; modification; infection; inflammation; tumour;		
XX	food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;		
KW	urinary tract infection; colitis; candida infection; fungicidal;		
KW	bactericidal; preservative.		
XX	Homo sapiens.		
OS	Synthetic.		
OS	WO200001730-A1.		
XX	13-JAN-2000.		
PD	06-JUL-1999;	99WO-SEQ01230.	
XX	06-JUL-1998;	98SE-0002441.	
PF	17-JUL-1998;	98SE-0002562.	
PR	29-DEC-1998;	98SE-0004614.	
PR	(ASCI-) A+ SCI INVEST AB.		
XX	Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;		
XX	WPI; 2000-147388/13.		
XX	New peptides used for treatment and prevention of infections,		
XX	inflammations and tumors and for use in infant formula food -		
XX	Claim 12; Page 69; 102pp; English.		
XX	AAV78001 to AAV78100 represent peptides having sequences based on human		
XX	lactoferrin. The peptides are taken up in the intestine through		
XX	binding to specific lactoferrin receptors and are then transported		
XX	through the circulation. A medicinal product of the peptide or fragment		
XX	can be used for treating and/or prevention of infections (such as		
XX	urinary tract infections, colitis, and Candida infection on a mucosal		
XX	membrane), inflammations and/or tumours. The peptides can also be used		
XX	in food stuffs such as infant formula food. The peptides are also		
XX	fungicidal and bactericidal and may also be used as preservatives.		
XX	Even though native human lactoferrin have been shown to have desired		
XX	anti-inflammatory anti-infectious and anti-tumoural properties they		
XX	cannot be used clinically on a broad basis because of high production		
XX	costs. Therefore, provision of peptides based on lactoferrin would		
XX	enable them to be used for the same purposes as lactoferrin at lower		
XX	cost.		
XX	Sequence	15 AA;	
XX	Query Match	97.0%; Score 64; DB 21; Length 15;	
XX	Best Local Similarity	91.7%; Pred. No. 9.8e-05;	
XX	Matches	11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	1	CFQWQRYMKRYR 12	
DB	4	CFQWQRYMKRYR 15	

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# 3M protein - protein search, using sw model

Run on: February 21, 2003, 07:50:40 ; Search time 8.93023 Seconds  
(without alignments)  
39.537 Million cell updates/sec

Title: US-09-743-107B-89  
Perfect score: 66  
Sequence: 1 CFQQRXMRKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pcp:  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pcp:  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pcp:  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pcp:  
5: /cgn2\_6/prodata/1/iaa/PCUTUS COMB.pcp:  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pcp:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

result No.	Score	Query Match	Length	ID	Description
1	64	97.0	18	1	US-08-204-487-3
2	64	97.0	18	2	US-08-465-948-8
3	64	97.0	18	2	US-08-628-380-8
4	64	97.0	18	2	US-08-475-055-8
5	64	97.0	20	1	US-07-755-161A-3
6	64	97.0	20	1	US-07-891-174-3
7	64	97.0	20	1	US-08-204-487-1
8	64	97.0	20	1	US-08-256-771-24
9	64	97.0	20	1	US-08-256-771-25
10	64	97.0	20	1	US-08-381-984-24
11	64	97.0	20	1	US-08-381-984-25
12	64	97.0	22	4	US-09-508-734-4
13	64	97.0	24	4	US-09-508-734-6
14	64	97.0	25	1	US-07-755-161A-10
15	64	97.0	25	1	US-07-891-174-10
16	64	97.0	25	1	US-08-204-487-7
17	64	97.0	29	4	US-09-508-734-8
18	64	97.0	36	1	US-07-755-161A-8
19	64	97.0	36	1	US-07-891-174-8
20	64	97.0	36	1	US-08-256-771-30
21	64	97.0	36	1	US-08-381-984-29
22	64	97.0	47	2	US-08-464-182A-6
23	64	97.0	47	2	US-08-406-271-6
24	64	97.0	50	2	US-08-693-274A-7
25	64	97.0	52	4	US-09-017-043A-3
26	64	97.0	53	2	US-08-464-182A-5
27	64	97.0	53	2	US-08-406-271-5

28	64	97.0	54	2	US-08-464-182A-2	Sequence 2, Appli
29	64	97.0	54	2	US-08-406-271-2	Sequence 2, Appli
30	64	97.0	594	3	US-08-724-586-2	Sequence 2, Appli
31	64	97.0	694	4	US-09-421-632-2	Sequence 2, Appli
32	64	97.0	694	4	US-09-932-190-2	Sequence 2, Appli
33	64	97.0	705	2	US-08-655-640-2	Sequence 2, Appli
34	64	97.0	708	2	US-08-655-640-4	Sequence 4, Appli
35	64	97.0	711	1	US-08-154-019-4	Sequence 4, Appli
36	64	97.0	711	1	US-08-461-333-4	Sequence 4, Appli
37	64	97.0	711	3	US-08-464-167-4	Sequence 4, Appli
38	64	97.0	711	3	US-09-158-313-4	Sequence 4, Appli
39	64	97.0	711	4	US-08-476-798-4	Sequence 4, Appli
40	61	92.4	711	1	US-08-145-308-2	Sequence 2, Appli
41	61	92.4	711	1	US-08-250-681-2	Sequence 2, Appli
42	61	92.4	711	1	US-08-453-703-2	Sequence 2, Appli
43	61	92.4	711	2	US-08-456-106-2	Sequence 2, Appli
44	61	92.4	711	3	US-08-456-108-2	Sequence 2, Appli
45	61	92.4	711	4	US-09-263-577-2	Sequence 2, Appli

## ALIGNMENTS

RESULT 1  
US-08-204-487-3  
; Sequence 3, Application US/08204487  
; Patent No. 5565425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAKOI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGEKI  
; APPLICANT: DOSAKO, SHUN'ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; TITLE OF INVENTION: INHIBITORS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESSEE: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FUN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..18  
; OTHER INFORMATION: /note= "HUMAN LACTOFERRIN PEPTIDE"



OTHER INFORMATION: (20-37)"  
US-08-204-487-3

Query Match 97.0%; Score 64; DB 1; Length 18;  
Best Local Similarity 91.7%; Pred. No. 3.8e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQXMKVR 12  
|||||  
Db 1 CFQWQXMKVR 12

## RESULT 2

US-08-485-948-8  
; Sequence 8, Application US/08485948  
; Patent No. 5855882

## GENERAL INFORMATION:

APPLICANT: YONG MING LI  
APPLICANT: HELEN VIASSARA  
APPLICANT: ANTHONY CERAMI  
TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/485,948

FILING DATE:

CLASSIFICATION: 436

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/488,217

FILING DATE: JUNE 7, 1995

APPLICATION NUMBER: 08/418,642

FILING DATE: APRIL 7, 1995

CLASSIFICATION: 436

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 947-1-008A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

DESCRIPTION: LF-Cl, 8-25

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

US-08-485-948-8

Query Match 97.0%; Score 64; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 3.8e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQXMKVR 12  
|||||  
Db 1 CFQWQXMKVR 12

## RESULT 3

US-08-628-380-8  
; Sequence 8, Application US/08628380  
; Patent No. 5891341

## GENERAL INFORMATION:

APPLICANT: LI, YONG MING  
APPLICANT: VIASSARA, HELEN  
APPLICANT: CERAMI, ANTHONY  
TITLE OF INVENTION: AGE-MEDIATED INHIBITION OF ANTIBACTERIAL PROTEINS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/628,380

FILING DATE: April 4, 1996

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/418,642

FILING DATE: APRIL 7, 1995

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 947-1-008 CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

DESCRIPTION: LF-Cl, 8-25

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

US-08-628-380-8

Query Match 97.0%; Score 64; DB 2; Length 18;

Best Local Similarity 91.7%; Pred. No. 3.8e-05;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQXMKVR 12

|||||

Db 1 CFQWQXMKVR 12

## RESULT 4

US-08-475-055-8  
; Sequence 8, Application US/08475055  
; Patent No. 5962245

## GENERAL INFORMATION:

APPLICANT: YONG MING LI  
APPLICANT: HELEN VIASSARA  
APPLICANT: ANTHONY CERAMI  
TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
NUMBER OF SEQUENCES: 9

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Klauber & Jackson  
;; STREET: 411 Hackensack Avenue  
;; CITY: Hackensack  
;; STATE: New Jersey  
;; COUNTRY: USA  
;; ZIP: 07601  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; OPERATING SYSTEM: IBM PC compatible  
;; SOFTWARE: Patent Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/475,055  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/485,948  
;; FILING DATE:  
;; APPLICATION NUMBER: 08/488,217  
;; FILING DATE: JUNE 7, 1995  
;; APPLICATION NUMBER: 08/418,642  
;; FILING DATE: APRIL 7, 1995  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Jackson Esq., David A.  
;; REGISTRATION NUMBER: 26,742  
;; REFERENCE/DOCKET NUMBER: 947-1-008A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 201 487-5800  
;; TELEFAX: 201 343-1684  
;; TELEX: 133521  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 18 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; DESCRIPTION: LF-CI, 8-25  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FRAGMENT TYPE: internal  
JS-08-475-055-8

Query Match 97.0%; Score 64; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred.No. 3.8e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2y 1 CFQWQXMKVR 12  
||| ||| |||  
2b 1 CFQWQXMKVR 12

RESULT 5  
JS-07-755-161A-3  
; Sequence 3, Application US/07755161A  
; Patent No. 5304633  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible

;; OPERATING SYSTEM: MS-DOS  
;; SOFTWARE: DisplayWrite  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/755,161A  
;; FILING DATE: 19910905  
;; CLASSIFICATION: 530  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Warren M. Cheek Jr.  
;; REGISTRATION NUMBER: 33,367  
;; REFERENCE/DOCKET NUMBER:  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-371-8850  
;; TELEFAX: 202-371-8856  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 20 amino acids  
;; TYPE: AMINO ACID  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE:  
;; HYPOTHETICAL:  
;; ANTI-SENSE:  
;; FRAGMENT TYPE:  
;; ORIGINAL SOURCE:  
;; ORGANISM:  
;; STRAIN:  
;; INDIVIDUAL ISOLATE:  
;; DEVELOPMENTAL STAGE:  
;; HAPLOTYPE:  
;; TISSUE TYPE:  
;; CELL TYPE:  
;; CELL LINE:  
;; ORGANELLE:  
;; IMMEDIATE SOURCE:  
;; LIBRARY:  
;; CLONE:  
;; POSITION IN GENOME:  
;; CHROMOSOME/SEGMENT:  
;; MAP POSITION:  
;; UNITS:  
;; FEATURE:  
;; NAME/KEY: modified site  
;; LOCATION: 2  
;; IDENTIFICATION METHOD:  
;; OTHER INFORMATION: /note= "thiol group of  
;; OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with  
;; OTHER INFORMATION: thiol group of Cys residue at location 19"  
;; FEATURE:  
;; NAME/KEY: modified site  
;; LOCATION: 19  
;; IDENTIFICATION METHOD:  
;; OTHER INFORMATION: /note= "thiol group of  
;; OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with  
;; OTHER INFORMATION: thiol group of Cys residue at location 2"  
;; PUBLICATION INFORMATION:  
;; AUTHORS:  
;; TITLE:  
;; JOURNAL:  
;; VOLUME:  
;; ISSUE:  
;; PAGES:  
;; DATE:  
;; DOCUMENT NUMBER:  
;; FILING DATE:  
;; PUBLICATION DATE:  
;; RELEVANT RESIDUES IN SEQ ID NO:  
US-07-755-161A-3

Query Match 97.0%; Score 64; DB 1; Length 20;

Best Local Similarity 91.7%; Pred. No. 4.2e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWQRMKVR 12  
Db 2 CFQWQRMKVR 13

RESULT 6  
US-07-891-174-3  
; Sequence 3, Application US/07891174  
; Patent No. 5317084  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: DisplayWrite  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/891,174  
; FILING DATE: 29-MAY-1992  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/755,161  
; FILING DATE: 05-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELEPHONE: 202-371-8850  
; TELEFAX: 202-371-8856  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; HYPOTHETICAL:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:  
; CELL LINE:  
; ORGANELLE:  
; IMMEDIATE SOURCE:  
; LIBRARY:  
; CLONE:  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT:  
; MAP POSITION:  
; UNITS:  
; FEATURE:  
; NAME/KEY: modified site

LOCATION: 2  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 19"  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 19  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 2"  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-891-174-3

Query Match 97.0%; Score 64; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4.2e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWQRMKVR 12  
Db 2 CFQWQRMKVR 13

RESULT 7  
US-08-204-487-1  
; Sequence 1, Application US/08204487  
; Patent No. 5565425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAOKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGEAKI  
; APPLICANT: DOSAKO, SHUN'ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; TITLE OF INVENTION: INHIBITORS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESSEE: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FJN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1-20  
OTHER INFORMATION: /note= "ANTIBACTERIAL PEPTIDE  
DERIVED FROM HUMAN LACTOFERRIN"  
US-08-204-487-1

Query Match 97.0%; Score 64; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4.2e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWQRXMRKVR 12  
| | | | | | | | | |  
Db 2 CFQWQRXMRKVR 13

RESULT 8  
US-08-256-771-24  
Sequence 24, Application US/08256771  
Patent No. 5656591  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
PRODUCTS THEREWITH  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,771  
FILING DATE: July 22, 1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "Cys residues are linked by  
OTHER INFORMATION: disulfide bond"

Query Match 97.0%; Score 64; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4.2e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWQRXMRKVR 12  
| | | | | | | | | |  
Db 2 CFQWQRXMRKVR 13

RESULT 10  
US-08-381-984-24

US-08-256-771-24

Query Match 97.0%; Score 64; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4.2e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWQRXMRKVR 12  
| | | | | | | | | |  
Db 2 CFQWQRXMRKVR 13

RESULT 9  
US-08-256-771-25  
Sequence 25, Application US/08256771  
Patent No. 5656591  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
PRODUCTS THEREWITH  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,771  
FILING DATE: July 22, 1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:

INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "Cys residues are protected to  
OTHER INFORMATION: prevent disulfide bond"

US-08-256-771-25

Query Match 97.0%; Score 64; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4.2e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWQRXMRKVR 12  
| | | | | | | | | |  
Db 2 CFQWQRXMRKVR 13

; Sequence 24, Application US/08381984  
; Patent No. 5804555  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: ANTIOXIDANT  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/381,984  
; FILING DATE: April 11, 1995  
; CLASSIFICATION: 252  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX:  
; TELEX:

## INFORMATION FOR SEQ ID NO: 24:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:

; NAME/KEY:  
; LOCATION:

; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "cysteine residues at positions 2  
; OTHER INFORMATION: and 19 are bonded by disulfide linkage"

; FEATURE:

; NAME/KEY:  
; LOCATION:

; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "the specified peptide as well as  
; OTHER INFORMATION: peptides including the specified peptide as a fragment thereof"

US-08-381-984-24

Query Match 97.0%; Score 64; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4.2e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12

DB 2 CFQWQRMKVR 13

## RESULT 11

US-08-381-984-25  
; Sequence 25, Application US/08381984  
; Patent No. 5804555  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: ANTIOXIDANT  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack

; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/381,984  
; FILING DATE: April 11, 1995  
; CLASSIFICATION: 252  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX:  
; TELEX:

## INFORMATION FOR SEQ ID NO: 25:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:

; NAME/KEY:  
; LOCATION:

; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "the specified peptide as well as  
; OTHER INFORMATION: peptides including the specified peptide as a fragment thereof"

; NAME/KEY:  
; LOCATION:

; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "cysteine residues at positions 2  
; OTHER INFORMATION: and 19 are chemically modified to prevent disulfide linkage"

US-08-381-984-25

Query Match 97.0%; Score 64; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4.2e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12

DB 2 CFQWQRMKVR 13

## RESULT 12

US-09-508-734-4  
; Sequence 4, Application US/09508734  
; Patent No. 6423509  
; GENERAL INFORMATION:  
; APPLICANT: Sanyang Genex Corporation  
; TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and  
; FILE REFERENCE: PA/SYG/00139  
; CURRENT APPLICATION NUMBER: US/09/508,734  
; CURRENT FILING DATE: 2000-06-01  
; PRIOR APPLICATION NUMBER: PCT/KE99/00373  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: KRI998-29351  
; PRIOR FILING DATE: 1998-07-13  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Kopatentin 1.71  
; SEQ ID NO 4

```
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
JS-09-508-734-4
Query Match          97.0%; Score 64; DB 4; Length 22;
Best Local Similarity 91.7%; Pred. No. 4.6e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2Y 1 CFQWQRMKVR 12
   |||||
Db 2 CFQWQRMKVR 13

RESULT 13
JS-09-508-734-6
; Sequence 6, Application US/09508734
; Patent No. 6423509
; GENERAL INFORMATION:
; APPLICANT: Samyang Genex Corporation
; TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and
; FILE REFERENCE: PA/SYG/00139
; CURRENT APPLICATION NUMBER: US/09/508,734
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/KR99/00373
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: KR1998-29351
; PRIOR FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: KopatentIn 1.71
; SEQ ID NO 6
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
JS-09-508-734-6

Query Match          97.0%; Score 64; DB 4; Length 24;
Best Local Similarity 91.7%; Pred. No. 5e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2Y 1 CFQWQRMKVR 12
   |||||
Db 3 CFQWQRMKVR 14

RESULT 14
JS-07-755-161A-10
; Sequence 10, Application US/07755161A
; Patent No. 5304633
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; TITLE OF INVENTION: Antimicrobial Agent
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Menderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/755,161A
; FILING DATE: 19910905
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
```

```
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 4
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 21"
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 21
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 4"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-07-755-161A-10

Query Match          97.0%; Score 64; DB 1; Length 25;
Best Local Similarity 91.7%; Pred. No. 5.2e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12
   |||||
Db 4 CFQWQRMKVR 15
```

RESULT 15  
US-07-891-174-10  
Sequence 10: Application US/07891174  
Patent No. 5317084  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: Antimicrobial Peptides and an  
TITLE OF INVENTION: Antimicrobial Agent  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/891,174  
FILING DATE: 29-MAY-1992  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/755,161  
FILING DATE: 05-SEP-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX: 202-371-8856  
TELEX:  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 4  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 21"  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 21

IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 4"  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-891-174-10  
Query Match 97.0%; Score 64; DB 1; Length 25;  
Best Local Similarity 91.7%; Pred. No. 5.2e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CFQWQKXMKVR 12  
| | | | | | | | | |  
Db 4 CFQWQKXMKVR 15  
Search completed: February 21, 2003, 08:04:26  
Job time : 9.93023 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 08:08:15 ; Search time 6.88372 Seconds  
(without alignments)  
54.162 Million cell updates/sec

Title: US-09-743-107b-89

Perfect score: 66

Sequence: 1 CFQWQRMKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 156504 seqs, 31069816 residues

Total number of hits satisfying chosen parameters: 156504

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	64	97.0	15	9	US-09-798-869-2
2	64	97.0	25	9	US-09-798-869-20
3	64	97.0	694	9	US-10-023-096-2
4	56	84.8	15	9	US-09-798-869-6
5	53	80.3	15	9	US-09-798-869-3
6	53	80.3	25	9	US-09-798-869-23
7	45	68.2	15	9	US-09-798-869-7
8	44	66.7	15	9	US-09-798-869-4
9	44	66.7	25	9	US-09-798-869-22
10	41	62.1	15	9	US-09-798-869-8
11	41	62.1	15	9	US-09-798-869-29
12	41	62.1	15	9	US-09-798-869-30
13	37	56.1	10	10	US-09-864-761-47985
14	37	56.1	489	9	US-09-888-320-2
15	36	54.5	209	10	US-09-904-536-8
16	36	54.5	209	10	US-09-904-536-9
17	36	54.5	209	10	US-09-904-536-11
18	36	54.5	209	10	US-09-904-536-12
19	36	54.5	209	10	US-09-904-536-13

Sequence 14, Appl  
Sequence 15, Appl  
Sequence 16, Appl  
Sequence 17, Appl  
Sequence 18, Appl  
Sequence 10, Appl  
Sequence 6, Appl  
Sequence 1, Appl  
Sequence 6, Appl  
Sequence 1, Appl  
Sequence 5715, Ap  
Sequence 1031, Ap  
Sequence 4, Appl  
Sequence 2, Appl  
Sequence 5, Appl  
Sequence 48879, A  
Sequence 2, Appl  
Sequence 44327, A  
Sequence 26, Appl  
Sequence 25, Appl  
Sequence 1, Appl  
Sequence 10, Appl  
Sequence 28, Appl  
Sequence 24, Appl  
Sequence 21, Appl  
Sequence 4, Appl

20 36 54.5 209 10 US-09-904-536-14  
21 36 54.5 209 10 US-09-904-536-15  
22 36 54.5 209 10 US-09-904-536-16  
23 36 54.5 209 10 US-09-904-536-17  
24 36 54.5 209 10 US-09-904-536-18  
25 36 54.5 212 10 US-09-904-536-10  
26 36 54.5 235 9 US-10-095-449-6  
27 36 54.5 235 10 US-09-448-378-1  
28 36 54.5 235 10 US-09-983-806-6  
29 36 54.5 235 10 US-09-904-536-1  
30 35 53.0 86 9 US-09-738-626-5715  
31 34 51.5 95 10 US-09-764-864-1031  
32 34 51.5 846 9 US-10-051-409-4  
33 34 51.5 1617 9 US-10-050-453A-2  
34 33 50.0 15 9 US-09-798-869-5  
35 33 50.0 46 10 US-09-864-761-48879  
36 33 50.0 579 10 US-09-824-588-2  
37 32.5 49.2 124 10 US-09-864-761-44327  
38 32 48.5 13 9 US-09-798-869-26  
39 32 48.5 14 9 US-09-798-869-25  
40 32 48.5 15 9 US-09-798-869-1  
41 32 48.5 15 9 US-09-798-869-10  
42 32 48.5 15 9 US-09-798-869-28  
43 32 48.5 18 9 US-09-798-869-24  
44 32 48.5 25 9 US-09-798-869-21  
45 32 48.5 25 10 US-09-909-652-4

#### ALIGNMENTS

RESULT 1  
US-09-798-869-2  
; Sequence 2, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-09-798-869-2  
Query Match 97.0%; Score 64; DB 9; Length 15;  
Best Local Similarity 91.7%; Pred. No. 2.5e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CFQWQRMKVR 12  
Db 3 CFQWQRMKVR 14  
RESULT 2  
US-09-798-869-20  
; Sequence 20, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON





Matches 8; Conservative 2; Mismatches 1; Indels 1; Gaps 0;

cy 1 CFQWQRMKRV 11  
|:|||||:  
db 3 CYQWQRMKRL 13

## RESULT 6

US-09-798-869-23

Sequence 23, Application US/09798869

Publication No. US20030022821A1

GENERAL INFORMATION:

APPLICANT: JOHN SIGURD SVENDSEN

APPLICANT: (YSTEIN REKDAL

APPLICANT: BALDUR SVEINEBJ (RNSSON

APPLICANT: LARS VORLAND

TITLE OF INVENTION: BIOACTIVE PEPTIDES

FILE REFERENCE: A34049-PCT-USA-A

CURRENT APPLICATION NUMBER: US/09/798,869

CURRENT FILING DATE: 2001-02-27

PRIOR APPLICATION NUMBER: PCT/GB99/02851

PRIOR FILING DATE: 1999-08-31

PRIOR APPLICATION NUMBER: GB9818938.4

PRIOR FILING DATE: 1998-08-28

NUMBER OF SEQ ID NOS: 30

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 23

LENGTH: 25

TYPE: PRT

ORGANISM: CAPRINE

US-09-798-869-23

Query Match

Best Local Similarity 80.3%; Score 53; DB 9; Length 25;

Matches 8; Conservative 2; Mismatches 1; Indels 1; Gaps 0;

cy 1 CFQWQRMKRV 11  
|:|||||:  
db 3 CYQWQRMKRL 13

## RESULT 7

US-09-798-869-7

Sequence 7, Application US/09798869

Publication No. US20030022821A1

GENERAL INFORMATION:

APPLICANT: JOHN SIGURD SVENDSEN

APPLICANT: (YSTEIN REKDAL

APPLICANT: BALDUR SVEINEBJ (RNSSON

APPLICANT: LARS VORLAND

TITLE OF INVENTION: BIOACTIVE PEPTIDES

FILE REFERENCE: A34049-PCT-USA-A

CURRENT APPLICATION NUMBER: US/09/798,869

CURRENT FILING DATE: 2001-02-27

PRIOR APPLICATION NUMBER: PCT/GB99/02851

PRIOR FILING DATE: 1999-08-31

PRIOR APPLICATION NUMBER: GB9818938.4

PRIOR FILING DATE: 1998-08-28

NUMBER OF SEQ ID NOS: 30

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7

LENGTH: 15

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: synthetic peptide (modified form of homo sapiens

OTHER INFORMATION: sequence)

S-09-798-869-7

Query Match

Best Local Similarity 58.2%; Score 45; DB 9; Length 15;

Matches 7; Conservative 2; Mismatches 2; Indels 2; Gaps 0;

cy 1 CFQWQRMKRV 11  
|:|||||:  
db 3 CYQWQRMKRL 13

## RESULT 8

US-09-798-869-4

Sequence 4, Application US/09798869

Publication No. US20030022821A1

GENERAL INFORMATION:

APPLICANT: JOHN SIGURD SVENDSEN

APPLICANT: (YSTEIN REKDAL

APPLICANT: BALDUR SVEINEBJ (RNSSON

APPLICANT: LARS VORLAND

TITLE OF INVENTION: BIOACTIVE PEPTIDES

FILE REFERENCE: A34049-PCT-USA-A

CURRENT APPLICATION NUMBER: US/09/798,869

CURRENT FILING DATE: 2001-02-27

PRIOR APPLICATION NUMBER: PCT/GB99/02851

PRIOR FILING DATE: 1999-08-31

PRIOR APPLICATION NUMBER: GB9818938.4

PRIOR FILING DATE: 1998-08-28

NUMBER OF SEQ ID NOS: 30

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 4

LENGTH: 15

TYPE: PRT

ORGANISM: MURINE

US-09-798-869-4

Query Match

Best Local Similarity 66.7%; Score 44; DB 9; Length 15;

Matches 7; Conservative 1; Mismatches 3; Indels 3; Gaps 0;

cy 1 CFQWQRMKRV 11  
|:|||||:  
db 3 CLRQWQRMKRV 13

## RESULT 9

US-09-798-869-22

Sequence 22, Application US/09798869

Publication No. US20030022821A1

GENERAL INFORMATION:

APPLICANT: JOHN SIGURD SVENDSEN

APPLICANT: (YSTEIN REKDAL

APPLICANT: BALDUR SVEINEBJ (RNSSON

APPLICANT: LARS VORLAND

TITLE OF INVENTION: BIOACTIVE PEPTIDES

FILE REFERENCE: A34049-PCT-USA-A

CURRENT APPLICATION NUMBER: US/09/798,869

CURRENT FILING DATE: 2001-02-27

PRIOR APPLICATION NUMBER: PCT/GB99/02851

PRIOR FILING DATE: 1999-08-31

PRIOR APPLICATION NUMBER: GB9818938.4

PRIOR FILING DATE: 1998-08-28

NUMBER OF SEQ ID NOS: 30

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 22

LENGTH: 25

TYPE: PRT

ORGANISM: MURINE

US-09-798-869-22

Query Match

Best Local Similarity 66.7%; Score 44; DB 9; Length 25;

Matches 7; Conservative 1; Mismatches 3; Indels 3; Gaps 0;

cy 1 CFQWQRMKRV 11  
|:|||||:  
db 3 CLRQWQRMKRV 13

RESULT 10  
US-09-798-869-8  
; Sequence 8, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of caprine  
; OTHER INFORMATION: sequence)  
US-09-798-869-8

Query Match 62.1%; Score 41; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.26;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQXRKRV 11  
| | | | |  
Db 3 CLRQWEMKVL 13

RESULT 11  
US-09-798-869-29  
; Sequence 29, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 29  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: BOVINE  
US-09-798-869-29

Query Match 62.1%; Score 41; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.26;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQXRKRV 11  
| | | | |  
Db 3 CFRQWEMKVL 13

RESULT 12  
US-09-798-869-30

; Sequence 30, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 30  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: BOVINE  
US-09-798-869-30

Query Match 62.1%; Score 41; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.26;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQXRKRV 11  
| | | | |  
Db 3 CFRQWEMKVL 13

RESULT 13  
US-09-864-761-47985  
; Sequence 47985, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Cherp, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecmca-X-1  
; CURRENT APPLICATION NUMBER: 2001-05-23  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/609,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 47985  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Homo sapiens

FEATURE:  
OTHER INFORMATION: MAP TO AL096701.14  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6  
OTHER INFORMATION: EST\_HUMAN HIT: AN294800.1, EVALUATE 1.00e-06  
JS-09-864-761-47985

Query Match 56.1%; Score 37; DB 10; Length 21;  
Best Local Similarity 83.3%; Pred. No. 1.8;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWR 6  
| | | | |  
Db 16 CFQWR 21

RESULT 14  
JS-09-888-320-2  
; Sequence 2, Application US/09888320  
; Publication No. US20030013090A1  
; GENERAL INFORMATION:  
; APPLICANT: Barry III, Clifton E.  
; APPLICANT: DeBarber, Andrea E.  
; APPLICANT: Mdiluli, Khisimuzi  
; APPLICANT: Bekker, Linda-Gail  
; APPLICANT: The Government of the United States of America  
; APPLICANT: as represented by The Secretary of the  
; APPLICANT: Department of Health and Human Services  
; TITLE OF INVENTION: Methods of Diagnosing Multidrug Resistant Tuberculosis  
; FILE REFERENCE: 015280-413100US  
; CURRENT APPLICATION NUMBER: US/09/888,320  
; CURRENT FILING DATE: 2001-06-22  
; PRIOR APPLICATION NUMBER: US 60/214,187  
; PRIOR FILING DATE: 2000-06-26  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 489  
; TYPE: PRT  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: wild-type EtAa monooxygenase (RV3854c, EthA)  
JS-09-888-320-2

Query Match 56.1%; Score 37; DB 9; Length 489;  
Best Local Similarity 54.5%; Pred. No. 35;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWRXMRKV 11  
| | | | |  
Db 253 CQWRPRMRKM 263

RESULT 15  
JS-09-904-536-8  
; Sequence 8, Application US/09904536  
; Patent No. US20020111475A1  
; GENERAL INFORMATION:

; APPLICANT: Graddis, Thomas J.  
; APPLICANT: McGrew, Jeffrey T.  
; TITLE OF INVENTION: FL73-L MUTANTS AND METHODS OF USE  
; FILE REFERENCE: 03260.0028  
; CURRENT APPLICATION NUMBER: US/09/904,536  
; CURRENT FILING DATE: 2001-07-16  
; PRIOR APPLICATION NUMBER: PRIOR APPLICATION: 09/109,100  
; PRIOR FILING DATE: 1999-07-02  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 209  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-904-536-8

Query Match 54.5%; Score 36; DB 10; Length 209;  
Best Local Similarity 50.0%; Pred. No. 23;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWRXMRK 10  
| | | | |  
Db 178 CLHWQTRRR 187

Search completed: February 21, 2003, 08:11:55  
Job time : 6.88372 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

DM protein - protein search, using sw model

Run on: February 21, 2003, 07:48:01 ; Search time 10,6047 Seconds  
(without alignments)  
108,784 Million cell updates/sec

Title: US-09-743-107B-89

Perfect score: 66

Sequence: 1 CFQWQRXKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 73.\*

1: PIR1.\*

2: PIR2.\*

3: PIR3.\*

4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

result	No.	Score	Query Match	Length	DB ID	Description
1	64	97.0	711	1	TFHUL	lactotransferrin precursor
2	53	80.3	708	2	JC2233	lactoferrin - goat
3	50	75.8	33	2	S52107	lactoferrin - sheep
4	44	66.7	707	1	A28438	lactoferrin precursor
5	43	65.2	4568	2	T08030	dysfibrinogen beta heavy
6	40	60.6	275	2	T22597	hypothetical prote
7	40	60.6	932	2	T28820	hypothetical prote
8	39	59.1	206	2	H97451	pyridoxamine 5'-ph
9	39	59.1	206	2	AB2670	pyridoxamine 5'-ph
10	39	59.1	208	2	AG3441	probable pyridoxam
11	38	57.6	289	2	G86403	33.3K hypothetical
12	38	57.6	515	2	T00510	probable cytochrom
13	38	57.6	543	2	T00513	cytochrome P450 ho
14	38	57.6	1135	2	T14803	phytochrome C - so
15	37	56.1	85	2	AG0794	polymyxin B resist
16	37	56.1	121	2	AH3147	hypothetical prote
17	37	56.1	255	2	B87515	signal peptidase I
18	37	56.1	433	2	T32605	hypothetical prote
19	37	56.1	489	2	C70655	probable monooxyge
20	37	56.1	584	2	C84325	hypothetical prote
21	36	54.5	235	2	I38440	flt3 ligand - huma
22	36	54.5	275	1	JC1113	interleukin-2 rece
23	36	54.5	275	1	S07442	interleukin-2 rece
24	36	54.5	456	2	C86624	hypothetical prote
25	36	54.5	456	2	H72000	hypothetical prote
26	36	54.5	1391	2	T20642	hypothetical prote
27	36	54.5	1397	2	B87988	protein F90C3.1 li
28	35	53.0	124	2	C96582	Flt3l.22 (imported
29	35	53.0	120	1	S36092	pyridoxamine-phosp

RESULT 1

TFHUL

lactotransferrin precursor [validated] - human

N;Alternate names: lactoferrin

C;Species: Homo sapiens (man)

C;Date: 31-Mar-1992 #sequence revision 21-Nov-1997 #text change 08-Dec-2000

C;Accession: G01394; S11228; A45401; S10324; S15953; S20841; S07160; A61169; A31000; S74

R;Cho. Y.

Submitted to the EMBL Data Library, March 1994

A;Reference number: G08820

A;Accession: G01394

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-711 <CHO>

A;Cross-references: EMBL:U07643; NID:G467236; PIDN:AB60324.1; PID:G467237

R;Rey, M.W.; Woloshuk, S.L.; deBoer, H.A.; Pieper, F.R.

Nucleic Acids Res. 18, 5288, 1990

A;Title: Complete nucleotide sequence of human mammary gland lactoferrin.

A;Reference number: S11228; MUID:90384839; PMID:2402455

A;Accession: S11228

A;Molecule type: mRNA

A;Residues: 1-148, 'T', 150-422, 'C', 424-711 <REV>

A;Cross-references: EMBL:X53961; NID:G34415; PIDN:CAA37914.1; PID:G34416

R;Teng, C.T.; Liu, Y.; Yang, N.; Walmer, D.; Panella, T.

Mol. Endocrinol. 6, 1969-1981, 1992

A;Title: Differential molecular mechanism of the estrogen action that regulates lactofer

A;Reference number: A45401; MUID:93125571; PMID:1480183

A;Accession: A45401

A;Molecule type: DNA

A;Residues: 1-15 <TEN>

A;Cross-references: GB:S52659; NID:G263311; PIDN:AB24877.1; PID:G263312

A;Experimental source: placenta

A;Note: sequence extracted from NCBI backbone (NCBIP:122202)

R;Powell, M.J.; Ogden, J.E.

Nucleic Acids Res. 18, 4013, 1990

A;Title: Nucleotide sequence of human lactoferrin cDNA.

A;Reference number: S10324; MUID:90326549; PMID:2374734

A;Accession: S10324

A;Molecule type: mRNA

A;Residues: 3-711 <POW>

A;Cross-references: EMBL:X52941; NID:G34411; PIDN:CAA37116.1; PID:G34412

R;Stowell, K.M.; Rado, T.A.; Funk, W.D.; Tweedie, J.W.

Biochem. J. 276, 349-355, 1991

A;Title: Expression of cloned human lactoferrin in baby-hamster kidney cells.

A;Reference number: S15853; MUID:91264786; PMID:2049066

A;Accession: S15853

A;Status: nucleic acid sequence not shown; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 20-31 <ST1>

A;Accession: S20841

A;Molecule type: Protein

A;Residues: 20-28, 'X', 30-31 <ST2>

R;Rado, T.A.; Wei, X.; Benz Jr., E.J.  
 Blood 70, 989-993, 1987  
 A;Title: Isolation of lactoferrin cDNA from a human myeloid library and expression of mRNA  
 A;Reference number: S07160; MUID:88001031; PMID:3477300  
 A;Accession: S07160  
 A;Molecule type: mRNA  
 A;Residues: 436-487, 'A', 489-711 <RAD>  
 A;Cross-references: EMBL:M18642; NID:g186815; PIDN:AAA66665.1; PID:g386855  
 R;Panella, T.J.; Liu, Y.; Huang, A.T.; Teng, C.T.  
 Cancer Res. 51, 3037-3043, 1991  
 A;Title: Polymorphism and altered methylation of the lactoferrin gene in normal leukocytes  
 A;Reference number: A61169; MUID:91235214; PMID:1674448  
 A;Accession: A61169  
 A;Status: not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 3701, SWKPVN' <PAN>  
 A;Experimental source: normal breast tissue  
 R;Metz-Boutigue, M.H.; Jolles, J.; Mazurier, J.; Schoentgen, F.; Legrand, D.; Spik, G.;  
 Eur. J. Biochem. 145, 659-666, 1984  
 A;Title: Human lactotransferrin: amino acid sequence and structural comparisons with other  
 A;Reference number: A31000; MUID:85076667; PMID:6510420  
 A;Accession: A31000  
 A;Molecule type: protein  
 A;Residues: 20-140,142-169,171-203,'L',205,'K',207-208,'K',210-385,'Q',387-391,'W',393-4  
 A;Note: this is the final paper in a series  
 R;Houen, G.; Hoegdall, E.V.; Barkholt, V.; Norskov, L.  
 Eur. J. Biochem. 241, 303-308, 1996  
 A;Title: Lactoferrin: similarity to diamine oxidase and purification by aminohexyl affinity  
 A;Reference number: S74119; MUID:97054624; PMID:8898921  
 A;Accession: S74119  
 A;Molecule type: protein  
 A;Residues: 'G',23-24,'R',26-27,'XX',30-32 <HOU>  
 A;Experimental source: neutrophil granulocytes  
 C;Genetics:  
 A;Gene: GDB:LTF  
 A;Cross-references: GDB:119368; OMIM:150210  
 A;Map position: 3q21-3q23  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein; iron binding; milk  
 F;1-19/Domain: signal sequence #status predicted <SIG>  
 F;20-711/Product: lactotransferrin #status experimental <MAT>  
 F;21-356/Domain: transferrin repeat homology <TRH1>  
 F;360-699/Domain: transferrin repeat homology <TRH2>  
 F;29-65,39-56,135-218,177-193,190-201,251-265,503-697,595-609/Disulfide bonds: #status e  
 F;357-498/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F;368-400,378-391,425-706,447-669,479-554,513-527,524-537,647-652/Disulfide bonds: #stat

Query Match 97.0%; Score 64; DB 1; Length 711;  
 Best Local Similarity 91.7%; Pred. No. 0.00083;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQXMRKV 12  
 | | | | | | | | | | | | | |  
 Db 39 CFQWQXMRKV 50

RESULT 2  
 JC2323  
 lactoferrin - goat  
 C;Species: Capra aegagrus hircus (domestic goat)  
 C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 07-May-1999  
 C;Accession: JC2323  
 R;Le Provost, F.; Nocart, M.; Guerin, G.; Martin, P.  
 Biochem. Biophys. Res. Commun. 203, 1324-1332, 1994  
 A;Title: Characterization of the goat lactoferrin cDNA: assignment of the relevant locus  
 A;Reference number: JC2323; MUID:94380047; PMID:8093048  
 A;Accession: JC2323  
 A;Molecule type: mRNA  
 A;Residues: 1-708 <LEP>  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; transferrin; glycoprotein  
 F;359-696/Domain: transferrin repeat homology <TRH2>  
 F;252,300,387,495,564/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 80.3%; Score 53; DB 2; Length 708;  
 Best Local Similarity 72.7%; Pred. No. 0.083;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQXMRKV 11  
 | | | | | | | | | | | | | |  
 Db 38 CYQWQXMRKV 48

RESULT 3  
 S52107  
 lactoferrin - sheep (fragment)  
 C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C;Date: 14-Jul-1995 #sequence\_revision 05-Dec-1997 #text\_change 05-Dec-1997  
 C;Accession: S52107  
 R;Qian, Z.Y.; Jolles, P.; Migliore-Samour, D.; Fiat, A.M.  
 Biochim. Biophys. Acta 1243, 25-32, 1995  
 A;Title: Isolation and characterization of sheep lactoferrin, an inhibitor of platelet  
 A;Reference number: S52107; MUID:95127729; PMID:7827104  
 A;Accession: S52107  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-33 <QIA>  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication

Query Match 75.8%; Score 50; DB 2; Length 33;  
 Best Local Similarity 63.6%; Pred. No. 0.016;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQXMRKV 11  
 | | | | | | | | | | | | | |  
 Db 19 CYQWQXMRKV 29

RESULT 4  
 A28438  
 lactoferrin precursor - mouse  
 N;Alternate names: lactotransferrin  
 C;Species: Mus musculus (house mouse)  
 C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C;Accession: A28438; A41205  
 R;Pentecost, B.T.; Teng, C.T.  
 J. Biol. Chem. 262, 10134-10139, 1987  
 A;Title: Lactotransferrin is the major estrogen inducible protein of mouse uterine secret  
 A;Reference number: A32596; MUID:87280033; PMID:3611056  
 A;Accession: A28438  
 A;Molecule type: mRNA  
 A;Residues: 3-707 <PEN>  
 A;Cross-references: EMBL:J03298  
 R;Liu, Y.; Teng, C.T.  
 J. Biol. Chem. 266, 21880-21885, 1991  
 A;Title: Characterization of estrogen-responsive mouse lactoferrin promoter.  
 A;Reference number: A41205; MUID:92042099; PMID:1939212  
 A;Accession: A41205  
 A;Molecule type: DNA  
 A;Residues: 1-15 <LIU>  
 A;Cross-references: GB:M74778  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein  
 F;1-19/Domain: signal sequence #status predicted <SIG>  
 F;20-707/Product: lactotransferrin #status predicted <MAT>  
 F;358-695/Domain: transferrin repeat homology <TRH2>  
 F;494/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 66.7%; Score 44; DB 1; Length 707;  
 Best Local Similarity 63.6%; Pred. No. 3.6;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQXMRKV 11  
 | | | | | | | | | | | | | |  
 Db 37 CLRQWQXMRKV 47

```

RESULT 5
y yin beta heavy chain - Chlamydomonas reinhardtii
;Species: Chlamydomonas reinhardtii
;Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 02-Feb-2001
;Accession: T08030
;Molecule type: DNA
;Cell Sci. 107, 635-644, 1994
;Title: Sequence analysis of the Chlamydomonas alpha and beta dynein heavy chain genes
;Reference number: Z16302; PMID:94274778; PMID:8006077
;Accession: T08030
;Status: translated from GB/EMBL/DBSJ
;Molecule type: DNA
;Residues: 1-4568 <MIT>
;Cross-references: EMBL:U02963; NID:G409965; PIDN:AAA19956.1; PID:G514215
;Experimental source: strain 21gr
;Genetics:
;Gene: ODA4
;Map position: IX
;Introns: 48/2; 74/3; 105/3; 159/1; 225/3; 358/3; 514/3; 729/2; 919/3; 1004/3; 1115/3;
3334/3; 3686/3; 3882/3; 4240/3
;Superfamily: dynein heavy chain, ciliary
;Keywords: nucleotide binding; P-loop
;1919-1928/Region: nucleotide-binding motif A (P-loop)
;2202-2209/Region: nucleotide-binding motif A (P-loop)
;2530-2537/Region: nucleotide-binding motif A (P-loop)

Query Match 65.2%; Score 43; DB 2; Length 4568;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

y 1 CFQQRXMRKVR 12
||||| : :
b 1852 CFQWSQLRVIQ 1863

RESULT 6
22597
y ypothetical protein F53H4.4 - Caenorhabditis elegans
;Species: Caenorhabditis elegans
;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
;Accession: T22597
;Dobson, R.
;Submitted to the EMBL Data Library, October 1996
;Reference number: Z19587
;Accession: T22597
;Status: preliminary; translated from GB/EMBL/DBSJ
;Molecule type: DNA
;Residues: 1-275 <WIL>
;Cross-references: EMBL:Z81089; PIDN:CAB03137.1; GSPDB:GNO00028; CESP:F53H4.4
;Experimental source: clone F53H4
;Genetics:
;Gene: CESP:F53H4.4
;Map position: X
;Introns: 67/1; 153/1
;Superfamily: Caenorhabditis elegans hypothetical protein F53H4.4

Query Match 60.8%; Score 40; DB 2; Length 275;
Best Local Similarity 63.6%; Pred. No. 7.8;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

y 2 FQWRXMRKVR 12
||||| : :
b 262 FQWISMRKTR 272

RESULT 7
28820
y ypothetical protein F07C3.1 - Caenorhabditis elegans
;Species: Caenorhabditis elegans
;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

```

```

;Accession: T28820
;Ravetto, A.; Gattung, S.
;Submitted to the EMBL Data Library, March 1996
;Description: the sequence of C. elegans cosmid F07C3.
;Reference number: Z20528
;Accession: T28820
;Status: preliminary; translated from GB/EMBL/DBSJ
;Molecule type: DNA
;Residues: 1-932 <FAV>
;Cross-references: EMBL:U50308; PIDN:AAC48001.1; GSPDB:GNO00023; CESP:F07C3.1
;Experimental source: strain Bristol N2; clone F07C3
;Genetics:
;Gene: CESP:F07C3.1
;Map position: 5
;Introns: 14/2; 55/3; 86/2; 159/3; 220/2; 268/3; 327/1; 363/3; 402/3; 463/2; 559/3; 599

Query Match 60.6%; Score 40; DB 2; Length 932;
Best Local Similarity 63.6%; Pred. No. 25;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 FQWRXMRKVR 12
||||| : :
DB 579 FQWRSARLVK 589

RESULT 8
H97451
y pyridoxamine 5'-phosphate oxidase (AF179611) [imported] - Agrobacterium tumefaciens (str
;Species: Agrobacterium tumefaciens
;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
;Accession: H97451
;R.Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
;A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum-
;Reference number: A97359; PMID:11743194
;Accession: H97451
;Status: preliminary
;Molecule type: DNA
;Residues: 1-206 <KUR>
;Cross-references: GB:AE007869; PIDN:AAK96569.1; PID:G15155733; GSPDB:GNO00169
;Genetics:
;Gene: AGR_C1381
;Map position: circular chromosome
;Superfamily: pyridoxamine-phosphate oxidase

Query Match 59.1%; Score 39; DB 2; Length 206;
Best Local Similarity 50.0%; Pred. No. 9;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWRXMRKVR 12
||||| : :
DB 88 CFHWKSLRRQVR 99

RESULT 9
AB2670
y pyridoxamine 5'-phosphate oxidase [imported] - Agrobacterium tumefaciens (strain C58, Du
;Species: Agrobacterium tumefaciens
;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
;Accession: AB2670
;R.Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
;Wood, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
;Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, J
ster, E.W.
;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
;Reference number: AB2577; PMID:11743193
;Accession: AB2670
;Status: preliminary
;Molecule type: DNA
;Residues: 1-206 <KUR>

```

A;Cross-references: GB:AE008688; PID:gl7739129; GSPDB:GN00186  
 A;Experimental source: strain C58 (Dupont)  
 C;Genetics:  
 A;Gene: puxH  
 C;Superfamily: pyridoxamine-phosphate oxidase

Query Match 59.1%; Score 39; DB 2; Length 206;  
 Best Local Similarity 50.0%; Pred. No. 9;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWQXMRKVR 12  
 |||:|:  
 Db 88 CFHWKSLRQVR 99

RESULT 10  
 33-3K hypothetical protein F28L5.13 - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 01-Feb-2002 #sequence\_revision 02-Mar-2001 #text\_change 09-Nov-2001  
 C;Accession: G86403  
 R;Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 C;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A;Reference number: A86141, MUID:21016719; PMID:11130712  
 A;Accession: G86403  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-289 <STO>  
 A;Cross-references: GB:AE005172; NID:gl0998925; PIDN:AAG26065.1; GSPDB:GN00141  
 C;Genetics:  
 A;Map position: 1

Query Match 59.1%; Score 39; DB 2; Length 208;  
 Best Local Similarity 50.0%; Pred. No. 9;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWQXMRKVR 12  
 |||:|:  
 Db 90 CFHWKSLRQVR 101

RESULT 11  
 33-3K hypothetical protein F28L5.13 - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Nov-2001  
 C;Accession: G86403  
 R;Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 C;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A;Reference number: A86141, MUID:21016719; PMID:11130712  
 A;Accession: G86403  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-289 <STO>  
 A;Cross-references: GB:AE005172; NID:gl0998925; PIDN:AAG26065.1; GSPDB:GN00141  
 C;Genetics:  
 A;Map position: 1

Query Match 57.6%; Score 38; DB 2; Length 289;  
 Best Local Similarity 50.0%; Pred. No. 19;  
 Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CFQWQXMRKVR 12  
 |||:|:  
 Db 8 CFTWEEYARHVR 19

RESULT 12  
 T00510  
 Probable cytochrome P450 At2g23220 [imported] - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 23-Mar-2001  
 C;Accession: T00510; A84622  
 R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul,  
 submitted to the EMBL Data Library, November 1997  
 A;Description: Arabidopsis thaliana chromosome II BAC T20D16 genomic sequence.  
 A;Reference number: Z14164  
 A;Accession: T00510  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-515 <ROU>  
 A;Cross-references: EMBL:AC002391; NID:G2642427; PID:G2642441  
 A;Experimental source: cultivar Columbia  
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A;Reference number: A84420; MUID:20083487; PMID:10617197  
 A;Accession: A84622  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-515 <STO>  
 A;Cross-references: GB:AE002093; NID:G2642441; PIDN:AA87109.1; GSPDB:GN00139  
 C;Genetics:  
 A;Gene: At2g23220; T20D16.15  
 A;Map position: 2  
 A;Introns: 182/3; 310/3  
 C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
 C;Keywords: heme; iron; metalloprotein  
 F;312-471/Domain: cytochrome P450 homology <P45>  
 F;449/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 57.6%; Score 38; DB 2; Length 515;  
 Best Local Similarity 83.3%; Pred. No. 33;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQXMRKVR 6  
 |||:|:  
 Db 469 CFQWQXMRKVR 474

RESULT 13  
 T00513  
 Cytochrome P450 homolog At2g23190 - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 23-Mar-2001  
 C;Accession: T00513; F84621  
 R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul,  
 submitted to the EMBL Data Library, November 1997  
 A;Description: Arabidopsis thaliana chromosome II BAC T20D16 genomic sequence.  
 A;Reference number: Z14164  
 A;Accession: T00513  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-543 <ROU>  
 A;Cross-references: EMBL:AC002391; NID:G2642427; PID:G2642444  
 A;Experimental source: cultivar Columbia  
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.



Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: F84621  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-543 <STO>  
A;Cross-references: GB:AE002093; NID:G2642444; PIDN:AB87112.1; GSPDB:GNC00139  
A;Genetics: 14803  
A;Gene: pmd  
A;Map position: 2  
A;Introns: 211/3; 339/3  
A;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
A;Keywords: heme; iron; metalloprotein  
A;341-500/Domain: cytochrome P450 homology <P45>  
A;478/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 57.6%; Score 38; DB 2; Length 543;  
Best Local Similarity 83.3%; Pred. No. 34;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRMKV 11  
| : | : | :  
b 498 CFQWQRMKV 503

RESULT 14  
14803  
phytochrome C - sorghum  
A;Species: Sorghum bicolor (sorghum)  
A;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 05-May-2000  
A;Accession: T14803  
A;Childs: K.L.; Miller, F.R.; Cordonnier-Pratt, M.M.; Pratt, L.H.; Morgan, P.W.; Mullet,  
submitted to the EMBL Data Library, April 1996  
A;Reference number: Z18186  
A;Accession: T14803  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-1135 <CHI>  
A;Cross-references: EMBL:U56731; NID:G1800218; PID:G1800219  
A;Genetics:  
A;Gene: PHYC  
A;Superfamily: phytochrome; phytochrome homology  
A;Keywords: chromoprotein; dimer; photoreceptor; phytochromobilin; transcription regulat  
A;65-581/Domain: phytochrome homology <PHY>  
A;321/Binding site: phytochromobilin (Cys) (covalent) #status predicted

Query Match 57.6%; Score 38; DB 2; Length 1135;  
Best Local Similarity 36.4%; Pred. No. 69;  
Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQRMKV 11  
| : | : | :  
b 775 CLEWKAQKI 785

RESULT 15  
AG0794  
polymyxin B resistance protein [imported] - *Salmonella enterica* subsp. *enterica* serovar  
A;Species: *Salmonella enterica* subsp. *enterica* serovar Typhi  
A;Note: this species has also been called *Salmonella typhi*  
A;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 09-Nov-2001  
A;Accession: AG0794  
A;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
h. T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
S.; Moule, S.; O'Garra, P.  
Nature 413, 848-852, 2001  
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.  
A;Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov  
A;Reference number: AB0502; PMID:11677608  
A;Accession: AG0794  
A;Status: preliminary  
A;Molecule type: DNA

A;Residues: 1-85 <PAR>  
A;Cross-references: GB:AL513382; PIDN:CAD07537.1; PID:G16503529; GSPDB:GN00176  
A;Genetics:  
A;Gene: pmd  
Query Match 56.1%; Score 37; DB 2; Length 85;  
Best Local Similarity 60.0%; Pred. No. 8.9;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQRMKV 10  
| : | : | :  
b 74 CDEWQLTRK 83

Search completed: February 21, 2003, 08:02:44  
Job time : 10.6047 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

M protein - protein search, using sw model

run on: February 21, 2003, 07:28:06 ; Search time 5.2093 Seconds  
(without alignments)  
95.544 Million cell updates/sec

title: US-09-743-107B-89

effect score: 66

sequence: 1 CFQWQKXKVR 12

coring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

searched: 112892 seqs, 41476328 residues

total number of hits satisfying chosen parameters: 112892

minimum DB seq length: 0

maximum DB seq length: 2000000000

post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

atabase : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

result No.	Score	Query Match	Length	ID	Description
1	64	97.0	711	1 TRFL_HUMAN	P02788 homo sapien
2	53	80.3	708	1 TRFL_CAMDR	Q9tun0 camelus dro
3	53	80.3	708	1 TRFL_CAPHI	Q29477 capra hircu
4	44	66.7	707	1 TRFL_MOUSE	P08071 mus musculu
5	43	65.2	4568	1 DYHB_CHIRE	Q39565 chlamydomon
6	39	59.1	695	1 TRFL_HORSE	O77811 equus cabal
7	38	57.6	292	1 NLA_DROME	Q9xz18 drosophila
8	38	57.6	1135	1 PHVC_SORBI	P93528 sorghum bic
9	37	56.1	85	1 PMRD_SALTY	P37589 salmonella
10	37	56.1	146	1 RPOB_LIBAF	P41187 liberibacte
11	36	54.5	235	1 FL3L_HUMAN	P49771 homo sapien
12	36	54.5	275	1 IL2A_BOVIN	P12342 bos taurus
13	36	54.5	275	1 IL2A_SHEEP	P26898 ovis aries
14	36	54.5	1137	1 PHVC_ORYSA	Q92w19 oryza sativ
15	36	54.5	1179	1 ATK1_ARATH	Q91t02 arabidopsis
16	35	53.0	270	1 PDHX_MYXXA	P21159 myxococcus
17	35	53.0	306	1 BUB2_YEAST	P26448 saccharomyc
18	35	53.0	435	1 DCOR_PANRE	P49725 panagrellus
19	35	53.0	502	1 C911_ARATH	Q9fg65 arabidopsis
20	35	53.0	783	1 YNR2_CAEEL	Q21988 caenorhabdi
21	35	53.0	2594	1 7LBS_DROVI	P20806 drosophila
22	34	51.5	211	1 LOLE_VIBCH	P57070 vibrio chol
23	34	51.5	227	1 CHEZ_PSEAE	Q51434 pseudomonas
24	34	51.5	275	1 VA16_VACCC	P16710 vaccinia vi
25	34	51.5	378	1 VA16_VACCC	P20993 vaccinia vi
26	34	51.5	455	1 RADA_CHLMU	Q9pk96 chlamydia m
27	34	51.5	500	1 TLCE_RICPR	O05962 rickettsia
28	34	51.5	708	1 TRFL_BUBBU	O77698 bubalus bub
29	34	51.5	785	1 Y008_HUMAN	Q15398 homo sapien
30	34	51.5	962	1 YBX7_SCHPO	Q10201 schizosacch
31	33	50.0	159	1 Y1GN_ECOLI	P31439 escherichia
32	33	50.0	272	1 IL2A_HUMAN	P01589 homo sapien
33	33	50.0	377	1 VA16_VARV	P33841 variola vir

## ALIGNMENTS

RESULT 1									
TRFL_HUMAN STANDARD; PRT; 711 AA.									
ID	P02788; Q16780; Q16785; Q16786; Q16789; Q00756; Q9H123; Q96KZ4;								Q9pj19 chlamydia m
AC	Q96KZ5;								P44693 haemophilus
DT	21-JUL-1986 (Rel. 01, Created)								Qute7 schizosacch
DT	15-JUL-1999 (Rel. 38, Last sequence update)								P50466 escherichia
DT	15-JUN-2002 (Rel. 41, Last annotation update)								P14632 sus scrofa
DE	Lactotransferrin precursor (Lactoferrin) [Contains: Lactoferrin A;								Q41045 pinus sylve
DE	Lactoferrin B; Lactoferrin C]								Q24325 drosophila
GN	LTF OR LP.								Q24325 drosophila
OS	Homo sapiens (Human).								Q4595 drosophila
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								P3368 drosophila
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								P24874 ascaris suu
OX	NCBI_TaxID=9606;								P31063 escherichia
RN	[1]								Q06399 salmonella
RP	SEQUENCE FROM N.A.								
RC	TISSUE=Mammary gland;								
XX	MEDLINE=90364339; PubMed=2402455;								
RA	Rey M.W., Woloshuk S.L., de Boer H.A., Pieper F.R.;								
RT	"Complete nucleotide sequence of human mammary gland lactoferrin.";								
RL	Nucleic Acids Res. 18:5288-5288(1990).								
RN	[2]								
RP	SEQUENCE FROM N.A.								
RC	TISSUE=Mammary gland;								
RA	Cho Y.Y.;								
RL	Thesis (1994), Genetic Engineering Research Institute / Taejon, Korea.								
RN	[3]								
RP	SEQUENCE FROM N.A.								
RA	Conneely O.M.;								
RT	Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.								
RL	[4]								
RN	SEQUENCE FROM N.A.								
RP	TISSUE=Mammary gland;								
RC	Liang Q., Jimenez-Flores R., Richardson T.;								
RA	"Molecular cloning and sequence analysis of human lactoferrin.";								
RT	Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.								
RN	[5]								
RP	SEQUENCE FROM N.A.								
RC	TISSUE=Bone marrow;								
RA	Wei X., Han J., Rado T.A.;								
RT	"Human neutrophil lactoferrin coding and 5' flanking region DNA								
RL	sequences.";								
RN	[6]								
RP	Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.								
RN	[7]								
RP	SEQUENCE FROM N.A.								
RC	TISSUE=Mammary gland;								
RA	Cheng H., Chen X., Huan L.;								
RT	"CDNA cloning and sequence analysis of human lactoferrin.";								
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.								
RN	[8]								
RP	SEQUENCE FROM N.A.								
RC	TISSUE=Prostate;								
RA	Strausberg R.;								
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.								
RN	[8]								

RP SEQUENCE OF 3-711 FROM N.A.  
RC TISSUE-Mammary gland;  
RX MEDLINE=90322649; PubMed=2374734;  
RA Powell M.J., Ogden J.E.,  
RL "Nucleotide sequence of human lactoferrin cDNA.";  
RN Nucleic Acids Res. 18:4013-4013(1990).  
[9]  
RP SEQUENCE OF 20-711.  
RX MEDLINE=85076667; PubMed=6510420;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Scheentgen F.,  
RA Legrand D., Spik G., Montreuil J., Jolles P.,  
RT "Human lactotransferrin: amino acid sequence and structural  
RT comparisons with other transferrins.";  
RL Eur. J. Biochem. 145:659-666(1984).  
[10]  
RP PRELIMINARY SEQUENCE OF 20-73; 134-171; 257-278 AND 347-530.  
RX MEDLINE=82046817; PubMed=6794640;  
RA Metz-Boutigue M.-H., Mazurier J., Jolles J., Spik G., Montreuil J.,  
RA Jolles P.,  
RT "The present state of the human lactotransferrin sequence. Study and  
RT alignment of the cyanogen bromide fragments and characterization of  
RT N- and C-terminal domains.";  
RN Biochim. Biophys. Acta 670:243-254(1981).  
[11]  
RP SEQUENCE OF 609-711.  
RX MEDLINE=82262043; PubMed=7049727;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,  
RA Jolles P.,  
RT "An 88 amino acid long C-terminal sequence of human  
RT lactotransferrin.";  
RL FEBS Lett. 142:107-110(1982).  
[12]  
RP SEQUENCE OF 436-711 FROM N.A.  
RX MEDLINE=88001031; PubMed=3477300;  
RA Rado T.A., Wei X., Benz E.J. Jr.,  
RT "Isolation of lactoferrin cDNA from a human myeloid library and  
RT expression of mRNA during normal and leukemic myelopoiesis.";  
RL Blood 70:989-993(1987).  
[13]  
RP SEQUENCE OF 237-711 FROM N.A.  
RA McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,  
RA Nhan M., Parnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,  
RA Gnoj L., la Bastide M., Kaplan N., Greco T., Touchman J., Muzny D.,  
RA Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M., Porcel B.M.,  
RA Dragan Y., Giscalone J., Pea A., Powell E., Solinsky K.A., Desilva U.,  
RA Diaz-Perez S., Zhou X., Yu Y., Watanabe M., Doggett N., Garcia D.,  
RA Sagripanti J.L.,  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
[14]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND REVISIONS.  
RX MEDLINE=90064528; PubMed=2585506;  
RA Anderson B.F., Baker H.M., Norris G.E., Rice D.W., Baker E.N.,  
RT "Structure of human lactoferrin: crystallographic structure analysis  
RT and refinement at 2.8-A resolution.";  
RL J. Mol. Biol. 209:711-734(1989).  
[15]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RA Haridas M., Anderson B.F., Baker E.N.,  
RT "Structure of human diferric lactoferrin refined at 2.2-A  
RT resolution.";  
RL Acta Crystallogr. D 51:629-646(1995).  
[16]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 20-353.  
RX MEDLINE=97156796; PubMed=9003186;  
RA Nicholson H., Anderson B.F., Eland T., Shewry S.C., Tweedie J.W.,  
RA Baker E.N.,  
RT "Mutagenesis of the histidine ligand in human lactoferrin: iron  
RT binding properties and crystal structure of the histidine-  
RT 253--methionine mutant.";  
RL Biochemistry 36:341-346(1997).  
[17]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX MEDLINE=99190892; PubMed=10089347;  
RA Sun X.L., Baker H.M., Shewry S.C., Jameson G.B., Baker E.N.,  
RT "Structure of recombinant human lactoferrin expressed in Aspergillus  
RT awamori.";  
RL Acta Crystallogr. D 55:403-407(1999).  
[18]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=99192677; PubMed=10089508;  
RA Jameson G.B., Anderson B.F., Norris G.E., Thomas D.H., Baker E.N.,  
RT "Structure of human apolactoferrin at 2.0-A resolution. Refinement  
RT and analysis of ligand-induced conformational change.";  
RL Acta Crystallogr. D 54:1319-1335(1998).  
[19]  
RP CHARACTERIZATION OF LACTOFERROXINS.  
RX MEDLINE=91166929; PubMed=1369293;  
RA Tani F., Iio K., Chiba H., Yoshikawa M.,  
RT "Isolation and characterization of opiod antagonist peptides derived  
RT from human lactoferrin.";  
RL Agric. Biol. Chem. 54:1803-1810(1990).  
[20]  
RP VARIANTS THR-30 AND ARG-48.  
RX PubMed=9873069;  
RA Klintworth G.K., Sommer J.R., Obrian G., Han L., Ahmed M.N.,  
RA Qumsiyeh M.B., Lin P.-Y., Basti S., Reddy M.K., Kanai A., Hotta Y.,  
RA Sugar J., Kumaramanickavel G., Munier F., Schorderet D.F.,  
RA El Matri L., Iwata F., Kaiser-Kupfer M., Nagata M., Nakayasu K.,  
RA Heitmanck J.F., Teng C.T.,  
RT "Familial subepithelial corneal amyloidosis (gelatinous drop-like  
RT corneal dystrophy): exclusion of linkage to lactoferrin gene.";  
RL Mol. Vision 4:31-32(1998).  
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
CC OF AN ANION, USUALLY BICARBONATE.  
CC -!- FUNCTION: LACTOFERROXINS A, B AND C HAVE OPIOID ANTAGONIST  
CC ACTIVITY. LACTOFERROXIN A SHOWS PREFERENCE FOR MU-RECEPTORS, WHILE  
CC LACTOFERROXIN B AND C HAVE SOMEWHAT HIGHER DEGREES OF PREFERENCE  
CC FOR KAPPA-RECEPTORS THAN FOR MU-RECEPTORS.  
CC -!- SUBUNIT: MONOMER.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
CC  
CC THE SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X53961; CAA37914.1; -  
CC EMBL; U07643; AAB60324.1; -  
CC EMBL; M93150; AAA36159.1; -  
CC EMBL; M83202; AAA59511.1; -  
CC EMBL; M83205; AAA58656.1; -  
CC EMBL; X18642; AAA86665.1; -  
CC EMBL; AF332168; AAG48753.1; -  
CC EMBL; BC015822; AAH15822.1; -  
CC EMBL; BC015823; AAH15823.1; -  
CC EMBL; M73700; AAA59479.1; -  
CC EMBL; X52941; CAA37116.1; -  
CC EMBL; U95626; AAB57795.1; -  
CC PIR; S11228; TFHUL.  
CC PDB; 1LFC; 31-AUG-94.  
CC PDB; 1LCF; 31-OCT-93.  
CC PDB; 1LFG; 31-JUL-94.  
CC PDB; 1LFH; 31-OCT-93.  
CC PDB; 1LFI; 31-OCT-93.  
CC PDB; 1LGB; 31-AUG-94.  
CC PDB; 1LGC; 31-AUG-94.  
CC PDB; 1BKA; 08-NOV-96.  
CC PDB; 1DSN; 08-MAR-96.  
CC PDB; 1HSE; 12-MAR-97.  
CC PDB; 1VFD; 21-APR-97.

FT	DISULFID	377	390	BY SIMILARITY.
FT	DISULFID	424	703	BY SIMILARITY.
FT	DISULFID	444	666	BY SIMILARITY.
FT	DISULFID	476	551	BY SIMILARITY.
FT	DISULFID	500	694	BY SIMILARITY.
FT	DISULFID	510	524	BY SIMILARITY.
FT	DISULFID	521	534	BY SIMILARITY.
FT	DISULFID	592	606	BY SIMILARITY.
FT	DISULFID	644	649	BY SIMILARITY.
FT	METAL	79	79	IRON 1 (BY SIMILARITY).
FT	METAL	111	111	IRON 1 (BY SIMILARITY).
FT	METAL	211	211	IRON 1 (BY SIMILARITY).
FT	METAL	272	272	IRON 1 (BY SIMILARITY).
FT	METAL	414	414	IRON 2 (BY SIMILARITY).
FT	METAL	452	452	IRON 2 (BY SIMILARITY).
FT	METAL	545	545	IRON 2 (BY SIMILARITY).
FT	METAL	614	614	IRON 2 (BY SIMILARITY).
FT	BINDING	140	140	ANION (BY SIMILARITY).
FT	BINDING	482	482	ANION (BY SIMILARITY).
FT	CARBOHYD	252	252	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	385	385	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	537	537	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	594	594	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	261	261	F -> S (IN REF. 2).
FT	CONFLICT	304	304	G -> A (IN REF. 2).
FT	CONFLICT	330	330	S -> P (IN REF. 2).
FT	CONFLICT	492	494	LIS -> PLF (IN REF. 2).
FT	CONFLICT	506	506	L -> F (IN REF. 2).
FT	CONFLICT	609	609	A -> P (IN REF. 2).
FT	CONFLICT	642	642	R -> Q (IN REF. 2).
FT	CONFLICT	708 AA;	77211 MW;	0B0C175A0B69D430 CRC64;
SEQ	SEQUENCE	80.3%;	Score 53;	DB 1; Length 708;
	Query Match	75.0%;	Pred. No. 0.025;	Indels 0; Gaps 0;
	Best Local Similarity	Conservative	1; Mismatches	2; Indels
QY	1 CFQWQRXMKVR 12			
DB	38 CAQWQRXMKVR 49			
RESULT 3				
TRFL	CAPHI	STANDARD;	PRT;	708 AA.
ID	TRFL CAPHI			
AC	Q29477; Q29479;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Lactotransferrin precursor (Lactoferrin).			
GN	LTP.			
OS	Capra hircus (Goat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Caprinae; Capra.			
OC	NCBI_TaxID=9925;			
RN	[1]			
RN	SEQUENCE FROM N.A.			
RP	TISSUE=Mammary gland;			
RC	Lee T. Yu S., Kim S., Lee K., Yu D.;			
RA	Submitted (APR-1996) to the EMBL/GenBank/DBSJ databases.			
RRL	[2]			
RN	SEQUENCE FROM N.A.			
RP	TISSUE=Mammary gland;			
RC	MEDLINE=94380047; PubMed=8093048;			
RX	le Provost F., Nocart M., Guerin G., Martin P.;			
RT	"Characterization of the goat lactoferrin cDNA. Assignment of the			
RT	relevant locus to bovine U12 syntenic group.";			
RT	Biochem. Biophys. Res. Commun. 203:1324-1332(1994).			
CC	-!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH			
CC	CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING			
CC	OF AN ANION, USUALLY CARBONATE.			
CC	-!- SUBUNIT: MONOMER (BY SIMILARITY).			
CC	-!- SUBCELLULAR LOCATION: Secreted.			





```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ010930; CAA09407.1; --
DR PDB; 1B1X; 02-DEC-98.
DR PDB; 1B7U; 02-FEB-99.
DR PDB; 1B7Z; 02-FEB-99.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 2.
DR PRINTS; PRO0422; TRANSFERRIN.
DR SMART; SM00094; TR FER; 2.
DR PROSITE; PS00205; TRANSFERRIN_1; 2.
DR PROSITE; PS00206; TRANSFERRIN_2; 2.
DR PROSITE; PS00207; TRANSFERRIN_3; 1.
DR Transprot; Iron transport; Glycoprotein; Metal-binding; Repeat;
KW .Signal; 3D-structure.
FT NON TER 1
FT SIGNAL <1 6
FT CHAIN 7 695 LACTOTRANSFERRIN.
FT REPEAT 7 350
FT DISULFID 15 51
FT DISULFID 25 42
FT DISULFID 121 204
FT DISULFID 163 179
FT DISULFID 166 189
FT DISULFID 176 187
FT DISULFID 237 251
FT DISULFID 354 386
FT DISULFID 364 377
FT DISULFID 411 650
FT DISULFID 431 653
FT DISULFID 463 538
FT DISULFID 487 681
FT DISULFID 497 511
FT DISULFID 508 521
FT DISULFID 579 593
FT DISULFID 631 636
FT METAL 66 66
FT METAL 98 98
FT METAL 198 198
FT METAL 259 259
FT METAL 401 401
FT METAL 439 439
FT METAL 532 532
FT METAL 601 601
FT METAL 127 127
FT BINDING 127 127
FT BINDING 143 143 (POTENTIAL).
FT CARBOHYD 287 287 (POTENTIAL).
FT CARBOHYD 482 482 (POTENTIAL).
FT CARBOHYD 482 482 (POTENTIAL).
SQ SEQUENCE 695 AA; 75991 MW; 07BB84D50E1B165D CRC64;

Query Match
Best Local Similarity 59.1%; Score 39; DB 1; Length 695;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQRXMKVR 12
Db 25 CAKFQRMKKVR 36

RESULT 7
NL DROME
ID NL DROME STANDARD; PRT; 292 AA.
AC Q9XZL8; Q9V391;
DT 30-MAY-2000 (Rel. 39, Created)

```

```

DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nebula protein.
GN NL A OR CG6072.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RN SEQUENCE FROM N.A.
RP McCormick A.V., Goldberg M.L.;
RA "Gene required for elongation of meiosis I spindle in Drosophila
RT females."
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Flockerzi A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -!- FUNCTION: REQUIRED FOR ELONGATION OF MEIOSIS I SPINDLE.
CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF147700; AAD33987.1; --
DR EMBL; AE03712; AAF55285.1; --
DR Flybase; FBgn0026629; nla.
SQ SEQUENCE 292 AA; 31423 MW; 64F1BBF5F68A6CF9 CRC64;

```

```

DR SMART; SMO0387; HATPaase C; 1.
DR SMART; SMO0388; HsKA; 1.
DR SMART; SMO0086; PAC; 1.
DR SMART; SMO0091; PAS; 2.
DR TIGRFAMs; TIGR00229; sensory_box; 2.
DR PROSITE; PS50109; HIS_KIN; 1.
DR PROSITE; PS50112; PAS; 2.
DR PROSITE; PS50245; PHYTOCHROME_1; FALSE_NEG.
DR PROSITE; PS50046; PHYTOCHROME_2; 1.
DR KW Transcription regulation; Photoreceptor; Phytochrome; Chromophore;
KW Repeat; Multigene family.
FT BINDING 321 321 CHROMOPHORE (BY SIMILARITY).
FT DOMAIN 618 688 PAS 1.
FT DOMAIN 748 822 PAS 2.
FT DOMAIN 902 1122 HISTIDINE KINASE.
FT BINDING 321 321 CHROMOPHORE (BY SIMILARITY).
SQ SEQUENCE 1135 AA; 126315 MW; AFCC934B7592D54D CRC64;

Query Match 57.6%; Score 38; DB 1; Length 1135;
Best Local Similarity 36.4%; Pred. No. 23;
Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQRXMRKV 11
| : : : :
DB 775 CLEWKNAMQKI 785

RESULT 9
PMRD_SALTY
ID PMRD SALTY STANDARD; PRT; 85 AA.
AC P37589;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Polymyxin B resistance protein pmrd.
GN PMRD-OR STM2304.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAINS-LT2;
RX MEDLINE=94266712; PubMed=8206837;
RA Roland K.L., Esther C.R., Spitznagel J.K.;
RT "Isolation and characterization of a gene, pmrd, from Salmonella
RT typhimurium that confers resistance to polymyxin when expressed in
RT multiple copies.";
RL J. Bacteriol. 176:3589-3597(1994).
[2]
RP SEQUENCE FROM N.A.
RC STRAINS-LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC -!- FUNCTION: CONFERS RESISTANCE TO POLYMYXIN B. POLYMYXIN RESISTANCE
CC MAY BE MEDIATED BY AN INTERACTION BETWEEN PMRA OR A PMRA-REGULATED
CC GENE PRODUCT AND PMRD.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch.
CC -----
CC EMBL; U02281; AAA21322.1; -

```



DR EMBL; AE008803; AAL21205.1; -;  
KW StyGene; SG10304; pmrD.  
KW Antibiotic resistance; Complete proteome.  
SQ SEQUENCE 85 AA; 9749 MW; 1E1822419EA500CA CRC64;  
  
Query Match 56.1%; Score 37; DB 1; Length 85;  
Best Local Similarity 60.0%; Pred. No. 2.7;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 CFQWQXMRK 10  
DB 74 CDEWQLTRK 83  
  
RESULT 10  
RPOB LIBAF STANDARD; PRT; 146 AA.  
AC P41187;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE DNA-directed RNA polymerase beta chain (EC 2.7.7.6) (Transcriptase  
DE beta chain) (RNA polymerase beta subunit) (Fragment).  
GN RPOB.  
OS Liberibacter africanus (Liberibacter africanum).  
OC Bacteria; Proteobacteria; Alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Candidatus Liberibacter.  
OX NCBI\_TaxID:34020;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Relaprat;  
RA Planet P., Jacoueix S., Bove J.M., Garnier M.;  
RT Detection and characterization of the African Citrus Greening  
RT Liberibacter by amplification, cloning and sequencing of the rplKJL-  
RT rpoBC operon.";  
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION  
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS  
CC SUBSTRATES.  
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
CC {RNA}(N).  
CC -!- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE  
CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1  
CC BETA' CHAIN.  
CC -!- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U09675; AAA19557.1; -;  
DR InterPro; IPR001572; RNA pol\_B.  
DR Pfam; PF00562; RNA pol B; 1.  
DR PROSITE; PS01166; RNA POL BETA; PARTIAL.  
KW Transferrase; Transcription; DNA-directed RNA polymerase.  
FT NON TER 146  
SQ SEQUENCE 146 AA; 16731 MW; 815D6EDD66FDB943 CRC64;  
  
Query Match 56.1%; Score 37; DB 1; Length 146;  
Best Local Similarity 60.0%; Pred. No. 4.6;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 CFQWQXMRK 10  
DB 10 CVQWSGARK 19  
  
RESULT 11  
FL3L HUMAN

ID ID FL3L HUMAN STANDARD; PRT; 235 AA.  
AC P49771;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE SL cytokine precursor (Fms-related tyrosine kinase 3 ligand) (Flt3  
DE ligand) (Flt3L).  
GN FLT3LG.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID:9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94195428; PubMed=8145851;  
RA Hannum C., Culpepper J., Campbell D., McClanahan T., Zurawski S.,  
RA Bazan J.F., Kastelein R., Hudak S., Wagner J., Mattson J., Luh J.,  
RA Duda G., Martina N., Peterson D., Menon S., Shanafelt A.,  
RA Muench M., Kelner G., Namikawa R., Rennick D., Roncarolo M.G.,  
RA Zlotnik A., Rosnet O., Dubreuil P., Birnbaum D., Lee F.;  
RT Ligand for Flt3/Flk2 receptor tyrosine kinase regulates growth of  
RT haematopoietic stem cells and is encoded by variant RNAs.";  
RL Nature 368:643-648(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94235842; PubMed=8180375;  
RA Lyman S.D., James L., Johnson L., Brasel K., de Vries P.,  
RA Escobar S.S., Downey H., Splet R.R., Beckmann M.P., McKenna H.J.;  
RT Cloning of the human homologue of the murine flt3 ligand: a growth  
RT factor for early hematopoietic progenitor cells.";  
RL Blood 83:2795-2801(1994).  
RN [3]  
RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.  
RX MEDLINE=96032581; PubMed=7566977;  
RA Lyman S.D., Stocking K., Davidson B., Fletcher F., Johnson L.,  
RA Escobar S.;  
RT "Structural analysis of human and murine flt3 ligand genomic loci.";  
RL Oncogene 11:1165-1172(1995).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX MEDLINE=20343011; PubMed=10881197;  
RA Savvides S.N., Boone T., Karplus P.A.;  
RT Flt3 ligand structure and unexpected commonalities of helical  
RT bundles and cystine knots.";  
RL Nat. Struct. Biol. 7:486-491(2000).  
CC -!- FUNCTION: STIMULATES THE PROLIFERATION OF EARLY HEMATOPOIETIC  
CC CELLS. SYNERGIZES WELL WITH A NUMBER OF OTHER COLONY STIMULATING  
CC FACTORS AND INTERLEUKINS.  
CC -!- SUBUNIT: Homodimer (isoform 2).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1);  
CC secreted (isoform 2).  
CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; a 1/membrane-bound (shown here)  
CC and 2/soluble; are produced by alternative splicing.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U04806; AAA17989.1; -;  
DR EMBL; U03858; AAA19825.1; -;  
DR EMBL; U29874; AAA90949.1; -;  
DR EMBL; U29874; AAA90950.1; -;  
DR PDB; 1ETE; 09-JUN-00.  
DR Genew; HGNC:3766; FLT3LG.  
DR MIM; 600007; -;  
DR InterPro; IPR004213; Flt3 lig.  
DR Pfam; PF02947; Flt3 lig; 1.  
KW Cytokine; Glycoprotein; Transmembrane; Alternative splicing; Signal;  
KW 3D-structure.

T SIGNAL 1 26  
T CHAIN 27 235  
T DOMAIN 27 184  
T TRANSEM 185 205  
T DOMAIN 206 235  
T DISULFID 30 111  
T DISULFID 70 153  
T DISULFID 119 158  
T CARBOHYD 126 126  
T CARBOHYD 149 149  
T VARSPLIC 161 178  
T VARSPLIC 179 235  
T CONFLICT 72 72  
T SEQUENCE 235 AA; 26416 MW; 73B95BF693B4CECF CRC64;  
Q SEQUENCE 235 AA; 26416 MW; 73B95BF693B4CECF CRC64;  
Query Match 54.5%; Score 36; DB 1; Length 235;  
Best Local Similarity 50.0%; Pred. No. 11;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Y 1 CFQWQXMRK 10  
b 204 CLWQRTERR 213

RESULT 12  
L2A BOVIN STANDARD; PRT; 275 AA.  
C P12342;  
T 01-OCT-1989 (Rel. 12, Created)  
T 01-OCT-1989 (Rel. 12, Last sequence update)  
T 16-OCT-2001 (Rel. 40, Last annotation update)  
E Interleukin-2 receptor alpha chain precursor (IL-2 receptor alpha subunit) (P55) (TAC antigen) (CD25).  
N IL2RA.  
S Bos taurus (Bovine).  
C Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
C Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
C Bovidae; Bovinae; Bos.  
X NCBI\_TaxID=9913;  
N [1]  
P SEQUENCE FROM N.A.  
X MEDLINE=88212503; PubMed=2835311;  
A Weinberg A.D., Shaw J., Faerkau V., Bleackley R.C., Magnuson N.S.,  
A Reeves R., Magnuson J.A.;  
T "Cloning of cDNA for the bovine IL-2 receptor (bovine Tac antigen).";  
L Immunology 63:603-610(1988).  
N [2]  
P SEQUENCE OF 1-21 FROM N.A.  
X MEDLINE=96116968; PubMed=85631178;  
A Yoo J., de Leon F.A., Stone R.T., Beattie C.W.;  
T "Cloning and chromosomal assignment of the bovine interleukin-2 receptor alpha (IL-2R alpha) gene";  
L Mamm. Genome 6:751-753(1995).

T 1- FUNCTION: RECEPTOR FOR INTERLEUKIN-2.  
C EXIST IN 3 DIFFERENT FORMS: A HIGH AFFINITY DIMER, AN INTERMEDIATE AFFINITY MONOMER (BETA CHAIN), AND A LOW AFFINITY MONOMER (ALPHA CHAIN). THE HIGH AND INTERMEDIATE AFFINITY FORMS ALSO ASSOCIATE WITH A GAMMA CHAIN.  
C 1- SUBCELLULAR LOCATION: Type I membrane protein.  
C 1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.  
C This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
R EMBL; M20818; AAA51414.1; --  
R EMBL; U24226; AAC48487.1; --

DR PIR; S07442; S07442.  
DR HSP; P01589; IILW.  
DR InterPro; IPR000436; Sushi\_SCR\_CCP.  
DR Pfam; PF00084; sushi; 2.  
DR SMART; SM00032; CCP; 2.  
KW Transmembrane; Glycoprotein; Receptor; Repeat; Signal; Sushi.  
FT SIGNAL 1 21  
FT CHAIN 22 275  
FT DOMAIN 22 243  
FT TRANSEM 244 252  
FT DOMAIN 263 275  
FT DOMAIN 23 78  
FT DOMAIN 122 185  
FT DISULFID 24 64  
FT DISULFID 251 77  
FT DISULFID 123 168  
FT DISULFID 152 184  
FT CARBOHYD 80 80  
FT CARBOHYD 109 109  
SQ SEQUENCE 275 AA; 31238 MW; 4901BBF9A4862390 CRC64;  
Query Match 54.5%; Score 36; DB 1; Length 275;  
Best Local Similarity 50.0%; Pred. No. 13;  
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CFQWQXMRKVR 12  
Db 261 CLWQKWKXNR 272

RESULT 13  
IL2A SHEEP STANDARD; PRT; 275 AA.  
AC P26896;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE Interleukin-2 receptor alpha chain precursor (IL-2 receptor alpha subunit) (P55) (TAC antigen) (CD25).  
GN IL2RA.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=T-cell;  
RA Verhagen A.A.;  
RN Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92241682; PubMed=1572550;  
RA Bujdosó R., Sargan D.R., Williamson M.L., McConnell I.;  
RT "Cloning of a cDNA encoding the ovine interleukin-2 receptor 55-kDa protein, CD25";  
RL Gene 113:283-284(1992).  
CC 1- FUNCTION: RECEPTOR FOR INTERLEUKIN-2.  
CC 1- SUBUNIT: NON COVALENT DIMER OF AN ALPHA AND A BETA CHAINS. IL-2R EXIST IN 3 DIFFERENT FORMS: A HIGH AFFINITY DIMER, AN INTERMEDIATE AFFINITY MONOMER (BETA CHAIN), AND A LOW AFFINITY MONOMER (ALPHA CHAIN). THE HIGH AND INTERMEDIATE AFFINITY FORMS ALSO ASSOCIATE WITH A GAMMA CHAIN.  
CC 1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC 1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
R EMBL; M20818; AAA51414.1; --  
R EMBL; U24226; AAC48487.1; --

```

CC EMBL; Z11560; CAA77652.1; -
CC EMBL; X60149; CAA42723.1; -
CC EMBL; A19167; CAA01447.1; -
CC PIR; S18910; S18910.
CC PIR; S18899; S18899.
CC PIR; J01113; J01113.
CC HSSP; P01589; I1LM.
CC InterPro; IPR000436; Sushi_SCR_CCP.
CC Pfam; PF00084; sushi; 2.
CC SMART; SM00032; CCP; 2.
KW Transmembrane; Glycoprotein; Receptor; Repeat; Signal; Sushi.
FT SIGNAL 1 21
FT CHAIN 22 275
FT DOMAIN 22 243
FT TRANSMEM 244 262
FT DOMAIN 263 275
FT DOMAIN 23 78
FT DOMAIN 122 185
FT DISULFID 24 64
FT DISULFID 751 77
FT DISULFID 123 168
FT DISULFID 152 184
FT CARBOHYD 80 80
FT CONFLICT 186 166
SQ SEQUENCE 275 AA; 30904 MW; 1101A2DE5ACA088 CRC64;

Query Match 54.5%; Score 36; DB 1; Length 275;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12
DB 261 CLTWQRWKKQR 272

RESULT 14
PHYC_ORYSA STANDARD; PRT; 1137 AA.
AC Q9ZMI9; P93429;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phytochrome C.
GN PHYC.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Tahir M., Kanegae H., Takano M.;
RT "Phytochrome C (PHYC) Gene in rice: isolation and characterization of
a complete coding sequence.";
RL (in) Plant Gene Register PCR98-210.
RN [2]
RP SEQUENCE OF 275-378 FROM N.A.
RX MEDLINE=97019052; PubMed=8865668;
RA Mathews S., Sharrock R.A.;
RT "The phytochrome gene family in grasses (Poaceae): a phylogeny and
evidence that grasses have a subset of the loci found in dicot
angiosperms.";
RL Mol. Biol. Evol. 13:1141-1150(1996).
CC -1- FUNCTION: REGULATORY PHOTORECEPTOR WHICH EXISTS IN TWO FORMS THAT
ARE REVERSIBLY INTERCONVERTIBLE BY LIGHT: THE PR FORM THAT ABSORBS
MAXIMALLY IN THE RED REGION OF THE SPECTRUM AND THE PFR FORM THAT
ABSORBS MAXIMALLY IN THE FAR-RED REGION. PHOTOCONVERSION OF PR IN
PFR INDUCES AN ARRAY OF MORPHOGENIC RESPONSES, WHEREAS
RECONVERSION OF PFR TO PR CANCELS THE INDUCTION OF THOSE
RESPONSES. PFR CONTROLS THE EXPRESSION OF A NUMBER OF NUCLEAR
GENES INCLUDING THOSE ENCODING THE SMALL SUBUNIT OF RIBULOSE-

```

```

CC BISPHOSPHATE CARBOXYLASE, CHLOROPHYLL A/B BINDING PROTEIN,
CC PROTOCHLOROPHYLLIDE REDUCTASE, RRNA, ETC. IT ALSO CONTROLS
CC THE EXPRESSION OF ITS OWN GENE(S) IN A NEGATIVE FEEDBACK FASHION.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- PTM: CONTAINS ONE COVALENTLY LINKED TETRAPYRROLE CHROMOPHORE.
CC -1- SIMILARITY: BELONGS TO THE PHYTOCHROME FAMILY.
CC -1- SIMILARITY: CONTAINS 2 PAS (PER-ARNT-SIM) DIMERIZATION DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch)
CC -----
CC EMBL; AB018442; BAA74448.1; -
CC EMBL; U61207; ABE41996.1; -
CC InterPro; IPR003594; ATPbind_ATPase.
CC InterPro; IPR003018; GAF.
CC InterPro; IPR004359; HIS_KIN_sig.
CC InterPro; IPR003661; His_KinA.
CC InterPro; IPR001610; PAC.
CC InterPro; IPR000014; PAS domain.
CC InterPro; IPR001294; Phytochrome.
CC Pfam; PF00360; phytochrome; 1.
CC Pfam; PF00512; signal; 1.
CC Pfam; PF00989; PAS; 2.
CC Pfam; PF01590; GAF; 1.
CC Pfam; PF02518; HATPase_C; 1.
CC PRINTS; PS01033; PHYTOCHROME.
CC SMART; SM00085; GAF; 1.
CC SMART; SM00387; HATPase_C; 1.
CC SMART; SM00388; HSKA; 1.
CC SMART; SM00086; PAC; 1.
CC SMART; SM00091; PAS; 2.
CC TIGRFAMs; TIGR00229; sensory_box; 2.
CC PROSITE; PS0109; HIS_KIN; 1.
CC PROSITE; PS0112; PAS; 2.
CC PROSITE; PS0245; PHYTOCHROME_1; 1.
CC PROSITE; PS0046; PHYTOCHROME_2; 1.
KW Transcription regulation; Photoreceptor; Phytochrome; Chromophore;
KW Repeat; Multigene family.
FT DOMAIN 620 690 PAS 1.
FT DOMAIN 750 824 PAS 2.
FT DOMAIN 904 1124 HISTIDINE KINASE.
FT BINDING 322 322 CHROMOPHORE (BY SIMILARITY).
FT CONFLICT 279 279 F -> S (IN REF. 2).
FT CONFLICT 292 292 C -> S (IN REF. 2).
SQ SEQUENCE 1137 AA; 125982 MW; F2A520181CFE7B32 CRC64;

Query Match 54.5%; Score 36; DB 1; Length 1137;
Best Local Similarity 36.4%; Pred. No. 54;
Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 11
DB 777 CLEWNEAMQKI 787

RESULT 15
ATX1_ARATH STANDARD; PRT; 1179 AA.
ID ATX1_ARATH
AC Q9LT02;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Potential cation-transporting ATPase (EC 3.6.3.-).
GN AT5G23630 OR MPM1.11.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

Db 1144 CYSWERLLR 1152  
Search completed: February 21, 2003, 07:51:37  
Job time : 6.2093 secs

NC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
XC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
XN NCBI\_TaxID=3702;  
XP [1]  
XQ SEQUENCE FROM N.A.  
XC STRAIN=cv. Columbia;  
XX MEDLINE=20181125; PubMed=10718197;  
YA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,  
IA Tabata S.;  
IT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence  
IT features of the regions of 3,076,755 bp covered by sixty P1 and TAC  
IT clones."  
XL DNA Res. 7:31-63 (2000).  
XC -I- FUNCTION: INVOLVED IN TRANSPORT OF CATIONS (POTENTIAL).  
XC -I- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.  
XC -I- SUBCELLULAR LOCATION: Integral membrane protein.  
XC -I- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY  
XC (E1-E2 ATPASES). SUBFAMILY V.  
XC -----  
XC This SWISS-PROT entry is copyright. It is produced through a collaboration  
XC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
XC the European Bioinformatics Institute. There are no restrictions on its  
XC use by non-profit institutions as long as its content is in no way  
XC modified and this statement is not removed. Usage by and for commercial  
XC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
XC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
XC -----  
XR EMBL; AB025633; BAA97238.1; -  
XR InterPro; IPR001757; ATPase\_E1-E2.  
XR InterPro; IPR001454; Hlgase/hydrilase.  
XR Pfam; PF00122; E1-E2\_ATPase; 1.  
XR Pfam; PF00702; Hydrolase; 1.  
XR PRINTS; PR00119; CATATPASE.  
XR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
CW Hydrolase; Transmembrane; Phosphorylation; ATP-binding; Metal-binding;  
CW Magnesium.  
TT DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).  
TT TRANSMEM 21 42 POTENTIAL.  
TT DOMAIN 43 50 EXTRACELLULAR (POTENTIAL).  
TT TRANSMEM 51 71 POTENTIAL..  
TT DOMAIN 72 192 CYTOPLASMIC (POTENTIAL).  
TT TRANSMEM 193 215 POTENTIAL.  
TT DOMAIN 216 218 EXTRACELLULAR (POTENTIAL).  
TT TRANSMEM 219 238 POTENTIAL.  
TT DOMAIN 239 402 CYTOPLASMIC (POTENTIAL).  
TT TRANSMEM 403 422 POTENTIAL.  
TT DOMAIN 423 435 EXTRACELLULAR (POTENTIAL).  
TT TRANSMEM 436 453 POTENTIAL..  
TT DOMAIN 454 947 CYTOPLASMIC (POTENTIAL).  
TT TRANSMEM 948 967 POTENTIAL.  
TT DOMAIN 968 979 EXTRACELLULAR (POTENTIAL).  
TT TRANSMEM 980 997 POTENTIAL.  
TT DOMAIN 998 1013 CYTOPLASMIC (POTENTIAL).  
TT TRANSMEM 1014 1034 POTENTIAL..  
TT DOMAIN 1035 1059 EXTRACELLULAR (POTENTIAL).  
TT TRANSMEM 1060 1079 POTENTIAL.  
TT DOMAIN 1080 1092 CYTOPLASMIC (POTENTIAL).  
TT TRANSMEM 1093 1110 POTENTIAL.  
TT DOMAIN 1111 1128 EXTRACELLULAR (POTENTIAL).  
TT TRANSMEM 1129 1148 POTENTIAL.  
TT DOMAIN 1149 1179 CYTOPLASMIC (POTENTIAL).  
TT TRANSMEM 1175 1178 POLY-DYS.  
TT DOMAIN 1175 1178 PHOSPHORYLATION (BY SIMILARITY).  
TT MOD RES 491 491 MAGNESIUM (BY SIMILARITY).  
TT METAL 812 812 MAGNESIUM (BY SIMILARITY).  
TT METAL 816 816 MAGNESIUM (BY SIMILARITY).  
XQ SEQUENCE 1179 AA; 131115 MW; 4A3E82D222A4D78 CRC64;

Query Match 54.5%; Score 36; DB 1; Length 1179;  
Best Local Similarity 44.4%; Pred.No. 56;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

NY 1 CFQWQXMR 9  
|:|:|:

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

3M protein - protein search, using sw model

Run on: February 21, 2003, 07:44:43 ; Search time 21.6744 Seconds  
(without alignments)  
114.078 Million cell updates/sec

Title: US-09-743-107B-89

Perfect score: 66

Sequence: 1 CFQWQRMXKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 21:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteriapi:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	89.4	711	4 Q8TCD2	Q8tcd2 homo sapien
2	55	83.3	38	4 Q9UCY5	Q9ucy5 homo sapien
3	50	75.8	33	6 Q9TR80	Q9tr80 ovis aries
4	41	62.1	105	10 Q9XFD5	Q9xfds oryza sativ
5	40	60.6	148	10 Q9XHP1	Q9xhpl sesamum ind
6	40	60.6	275	5 Q93780	Q93780 caenorhabdi
7	40	60.6	932	5 Q19153	Q19153 caenorhabdi
8	39	59.1	205	16 Q986A0	Q986a0 rhizobium l
9	39	59.1	206	16 Q8UHC2	Q8unc2 agrobacteri
10	39	59.1	206	16 Q92RH8	Q92rh8 rhizobium m
11	39	59.1	208	16 Q8YFK3	Q8yfk3 brucella me
12	39	59.1	279	16 Q8XSE2	Q8xse2 ralstonia s
13	39	59.1	306	4 Q8TAX2	Q8tax2 homo sapien
14	39	59.1	466	4 Q9NU52	Q9nu52 homo sapien
15	38	57.6	91	15 Q77855	Q77855 human immun
16	38	57.6	91	15 Q77856	Q77856 human immun

17	38	57.6	274	4 Q96W21	Q96w21 homo sapien
18	38	57.6	289	10 Q9C6N2	Q9c6n2 arabidopsis
19	38	57.6	515	10 Q218S5	Q218s5 arabidopsis
20	38	57.6	543	10 Q21889	Q21889 arabidopsis
21	38	57.6	1121	10 Q9SWS6	Q9sws6 lycopersico
22	37	56.1	85	16 Q8Z536	Q8z536 salmonella
23	37	56.1	109	15 Q9YQC1	Q9yqci human immun
24	37	56.1	109	15 Q9YQCO	Q9yqco human immun
25	37	56.1	109	15 Q9YQ89	Q9yqb3 human immun
26	37	56.1	109	15 Q9YQ88	Q9yqb3 human immun
27	37	56.1	109	15 Q9YJ17	Q9yj17 human immun
28	37	56.1	109	15 Q9YJ12	Q9yj12 human immun
29	37	56.1	115	15 Q9Q9L0	Q9q9l0 human immun
30	37	56.1	119	15 Q8Q454	Q8q454 human immun
31	37	56.1	121	16 Q8U6K3	Q8u6k3 agrobacteri
32	37	56.1	255	16 Q9A6E4	Q9a6e4 caulobacter
33	37	56.1	433	5 Q44473	Q44473 caenorhabdi
34	37	56.1	459	4 Q9NZW0	Q9nzw0 homo sapien
35	37	56.1	460	4 Q9NZW3	Q9nzw3 homo sapien
36	37	56.1	489	16 P96223	P96223 mycobacteri
37	37	56.1	509	10 Q9LHAL	Q9lhal arabidopsis
38	37	56.1	584	17 Q9HPA3	Q9hpa3 halobacteri
39	37	56.1	632	4 Q94937	Q94937 homo sapien
40	37	56.1	864	5 Q62582	Q62582 encephalito
41	37	56.1	864	5 Q8SRG3	Q8srg3 encephalito
42	37	56.1	864	5 Q89Q16	Q8sq16 encephalito
43	37	56.1	866	10 Q9FHI9	Q9fhi9 arabidopsis
44	36	54.5	377	12 Q91MQ5	Q91mq5 lumphy skin
45	36	54.5	393	10 Q9ZTP0	Q9ztp0 oryza sativ

## ALIGNMENTS

RESULT 1

Q8TCD2 PRELIMINARY; PRT; 711 AA.

AC Q8TCD2; TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DE Lactotransferrin.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PROSTATE;  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC022347; AAH22347.1; -;  
SQ SEQUENCE 711 AA; 78327 MW; 1B9C7E097C45FAF CRC64;

Query Match 89.4%; Score 59; DB 4; Length 711;

Best Local Similarity 90.9%; Pred. No. 0.0067; Mismatches 0; Indels 0; Gaps 0;

Matches 10; Conservative

Oy 1 CFQWQRMXKVR 11

|||||

Db 39 CFQWQRMXKVR 49

|||||

RESULT 2

Q9UCY5 PRELIMINARY; PRT; 38 AA.

ID Q9UCY5; TREMBLrel. 13, Created)

AC Q9UCY5; TREMBLrel. 13, Last sequence update)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DE Lactoferrin homolog (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]\_SEQUENCE FROM N.A.  
RX MEDLINE=96081613; PubMed=8551695;  
RA Sato I.;  
RT "Characterization of the 84-kDa protein with ABH activity in human  
RT seminal plasma.";  
RL Jpn. J. Legal Med. 49:281-293 (1995).  
DR HSP; P02788; 18KA.  
DR InterPro; IPR001156; Transferrin.  
DR Pfam; PF00405; transferrin; 1.  
SQ SEQUENCE 38 AA; 4459 MW; 0402F490B5EBDEB CRC64;

Query Match 83.3%; Score 55; DB 4; Length 38;  
Best Local Similarity 90.9%; Pred. No. 0.0019;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FQWQXMRKVR 12  
DB 21 FQWQXMRKVR 31

## RESULT 3

Q9TR80 ID Q9TR80 PRELIMINARY; PRT; 33 AA.  
AC Q9TR80;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE Lactoferrin (Fragment).  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]\_SEQUENCE.  
RP MEDLINE=95127729; PubMed=7827104;  
RA Qian Z.Y., Jolles P., Migliore-Samouri D., Fiat A.M.;  
RL Biochim. Biophys. Acta 1243:25-32 (1995).  
DR HSP; O77698; 1CE2.  
DR InterPro; IPR001156; Transferrin.  
DR Pfam; PF00405; transferrin; 1.  
SQ SEQUENCE 33 AA; 3914 MW; D1904CAB15A73961 CRC64;

Query Match 75.8%; Score 50; DB 6; Length 33;  
Best Local Similarity 63.6%; Pred. No. 0.014;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQXMRKV 11  
DB 19 CFQWQXMRKV 29

## RESULT 4

Q9XFD5 ID Q9XFD5 PRELIMINARY; PRT; 105 AA.  
AC Q9XFD5;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Cytochrome P450 (Fragment).  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]\_SEQUENCE FROM N.A.  
RP TISSUE=PANICLE;  
RA Liu J., Yang J.;  
RT "Suppression subtractive hybridization (SSH) identified candidate

RT genes that are differentially expressed at rice young panicle.";  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
CC [1]\_SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
DR EMBL; AF140486; AAD29899.1; Cytochrome\_P450.  
DR InterPro; IPR001128; Cytochrome\_P450.  
DR Pfam; PF00067; P450; 1.  
DR PRINTS; PR00385; P450.  
DR PROSITE; PS00086; CYTOCHROME\_P450; UNKNOWN\_1.  
KW Heme; Monooxygenase; Oxidoreductase.  
FT NON TER 1  
SQ SEQUENCE 105 AA; 11912 MW; B0E3FCD487E19F9 CRC64;

Query Match 62.1%; Score 41; DB 10; Length 105;  
Best Local Similarity 60.0%; Pred. No. 2.3;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQXMRK 10  
DB 61 CFQWQXMRK 70

## RESULT 5

Q9XHP1 ID Q9XHP1 PRELIMINARY; PRT; 148 AA.  
AC Q9XHP1;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE 2S albumin.  
OS Sesamum indicum (Oriental sesame) (Gingelly).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Lamiales; Pedaliaceae; Sesamum.  
OX NCBI\_TaxID=4182;  
RN [1]\_SEQUENCE FROM N.A.  
RP STRAIN=TAINAN 1;  
RC MEDLINE=20074970; PubMed=10606554;  
RA Tai S.S.K., Wu L.S.H., Chen E.C.P., Tzen J.T.C.;  
RT "Molecular cloning of 11S globulin and 2S albumin, the two major seed  
RT storage proteins in sesame.";  
RL J. Agric. Food Chem. 47:4932-4938 (1999).  
DR EMBL; AF091841; AAD42943.1; -.  
DR InterPro; IPR003612; AAI.  
DR InterPro; IPR000617; Napin.  
DR InterPro; IPR001768; Try/amyL\_inhbr.  
DR Pfam; PF00234; tryp\_alpha\_amyL; 1.  
DR PRINTS; PR00496; NAPIN.  
DR SMART; SMO0499; AAI; 1.  
SQ SEQUENCE 148 AA; 17524 MW; BA46B033BA13E3DE CRC64;

Query Match 60.6%; Score 40; DB 10; Length 148;  
Best Local Similarity 66.7%; Pred. No. 5.1;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQXMRK 9  
DB 54 CFQWQXMRK 62

## RESULT 6

Q93780 ID Q93780 PRELIMINARY; PRT; 275 AA.  
AC Q93780;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE F53H4.4 protein.  
GN F53H4.4.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;

```

N [1]
P SEQUENCE FROM N.A.
L Dobson R.; (OCT-1996) to the EMBL/GenBank/DBJ databases.
L [2]
N Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
P P SEQUENCE FROM N.A.
X MEDLINE=99069613; PubMed=9851916;
A none;
T "Genome sequence of the nematode C.elegans: A platform for
T investigating biology.";
L Science 282:2012-2018(1998).
R ENBL; Z81089; CAB03137.1; -.
Q SEQUENCE 275 AA; 31716 MW; 9D69524FFC704DB2 CRC64;

Query Match 60.6%; Score 40; DB 5; Length 275;
Best Local Similarity 63.6%; Pred. No. 9.7;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Y 2 FQWQXMRKVR 12
b 262 FQWISMRKTR 272
|||||
|||||

RESULT 7
19153 PRELIMINARY; PRT; 932 AA.
D Q19153
C Q19153;
T 01-NOV-1996 (TrEMBLrel. 01, Created)
T 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
T 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
E Hypothetical 105.1 kDa protein.
N F07C3.1.
S Caenorhabditis elegans.
C Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
C Rhabditidae; Peloderinae; Caenorhabditis.
X NCBI_TaxID=6239;
N [1]
P SEQUENCE FROM N.A.
C STRAIN=BRISTOL N2;
X MEDLINE=99069613; PubMed=9851916;
A None;
T "Genome sequence of the nematode C. elegans: a platform for
T investigating biology. The C. elegans Sequencing Consortium.";
L Science 282:2012-2018(1998).
N [2]
P SEQUENCE FROM N.A.
C STRAIN=BRISTOL N2;
A Pavello A.; Gattung S.;
T "The sequence of C. elegans cosmid F07C3.";
L Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
N [3]
P SEQUENCE FROM N.A.
C STRAIN=BRISTOL N2;
A Waterston R.;
T "Direct Submission.";
L Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
R ENBL; U50308; AAG24025.1; -.
R InterPro; IPR000731; HMGR/patch_5TM.
R PROSITE; PS00156; SSD; 1.
W Hypothetical protein.
Q SEQUENCE 932 AA; 105144 MW; 66680619ADACBFD5 CRC64;

Query Match 60.6%; Score 40; DB 5; Length 932;
Best Local Similarity 63.6%; Pred. No. 34;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Y 2 FQWQXMRKVR 12
b 579 FQWQSRRLVK 589
|||||
|||||

RESULT 8
986A0 PRELIMINARY; PRT; 205 AA.
Q986A0;
AC Q986A0;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Pyridoxamine 5'-phosphate oxidase.
GN MLL7454.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VAFE303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Katanabe A., Ideasa Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Ideasa Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP004011; BAB5353.1; -.
DR InterPro; IPR000659; Pyridox_oxidase.
DR Pfam; PF01243; Pyridox_oxidase; 1.
DR ProDom; PD006312; Pyridox_oxidase; 1.
DR TIGRFAMs; TIGR00558; PdxH; 1.
DR PROSITE; PS01064; PYRIDOX_OXIDASE; 1.
KW Complete proteome.
SQ SEQUENCE 205 AA; 23300 MW; 0BADE4CD12327EA CRC64;

Query Match 59.1%; Score 39; DB 16; Length 205;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWQXMRKVR 12
Db 88 CFHWKSLRQVR 99
|||||
|||||

RESULT 9
Q8UHC2 PRELIMINARY; PRT; 206 AA.
ID Q8UHC2;
AC Q8UHC2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Pyridoxamine 5'-phosphate oxidase.
GN PDXH OR ATU0760 OR AGR_C1381.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RX SEQUENCE FROM N.A.

```





Query Match 59.1%; Score 39; DB 4; Length 466;  
 Best Local Similarity 55.6%; Pred. No. 26;  
 Matches 5; Conservative 2; Mismatches 0; Gaps 0;

QY 1 CFQWQXMR 9  
 ||||: :  
 Db 269 CFQWESTLR 277

## RESULT 15

Q77855 PRELIMINARY; PRT; 91 AA.  
 AC Q77855  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Envelope protein (Fragment).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OC NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95191002; PubMed=7884875;  
 RA Mulder-Kampinea G.A., Simonon A., Kuiken C.L., Dekker J.,  
 RA Scherpbier H.J., de Perre P., Boer K., Goudsmid J.,  
 RT "Similarity in env and gag genes between genomic RNAs of human  
 RT immunodeficiency virus type 1 (HIV-1) from mother and infant is  
 RL J. Virol. 69:2285-2296(1995).  
 DR EMBL; Z47867; CAA87881.1; -.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam; PF00516; GP120.1.  
 KW AIDS; Coat protein; Glycoprotein.  
 FT NON\_TER 1  
 FT TER 91  
 SQ SEQUENCE 91 AA; 10530 MW; 8B10C62011F305D6 CRC64;

Query Match 57.6%; Score 38; DB 15; Length 91;  
 Best Local Similarity 60.0%; Pred. No. 7.4;  
 Matches 6; Conservative 2; Mismatches 0; Gaps 0;

QY 3 QWQXMRKVR 12  
 ||||: :  
 Db 67 QWQXMRKVR 76

Search completed: February 21, 2003, 08:00:43  
 Job time : 21.6744 secs

Y 1 CFQWQXMRKVR 12  
 ||||: :  
 b 244 CFQWQXMRKVR 255

## RESULT 13

Q8TAX2 PRELIMINARY; PRT; 306 AA.  
 AC Q8TAX2  
 DT 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Similar to hypothetical protein FL011175.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 NC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 OC TISSUE=PANCREAS;  
 A Strausberg R.;  
 L Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 R EMBL; BC025708; AAH25708.1; -.  
 W Hypothetical protein.  
 Q SEQUENCE 306 AA; 34551 MW; 43D9AD722566B343 CRC64;

Query Match 59.1%; Score 39; DB 4; Length 306;  
 Best Local Similarity 55.6%; Pred. No. 17;  
 Matches 5; Conservative 2; Mismatches 0; Gaps 0;

Y 1 CFQWQXMR 9  
 ||||: :  
 b 269 CFQWESTLR 277

## RESULT 14

Q8NUS2 PRELIMINARY; PRT; 466 AA.  
 AC Q8NUS2  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 T CDNA FLJ1175 fis, clone PLACE1007375, weakly similar to phorbol  
 E ester/diacylglycerol-binding protein UNC-13.  
 S Homo sapiens (Human).  
 C Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 C Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 X NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 C TISSUE=PLACENTA;  
 A Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 A Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
 A Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 A Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
 A Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
 A Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y.,  
 A Ninomiya K., Iwayanagi T.;  
 T "NEDO human cDNA sequencing project."  
 L Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 R EMBL; AK002037; BA92048.1; -.  
 R HSP; P21707; IBYN.  
 R InterPro: IPR000008; C2.  
 R InterPro: IPR000504; RNA\_rec\_mot.  
 R Pfam; PF00168; C2; 1.  
 R PRINTS; PR00360; C2DOMAIN.  
 R SMART; SM0239; C2; 1.  
 R PROSITE; PS00499; C2\_DOMAIN\_1; UNKNOWN\_1.  
 R PROSITE; PS00004; C2\_DOMAIN\_2; 1.  
 R PROSITE; PS00030; RRM\_RNP\_1; UNKNOWN\_1.  
 Q SEQUENCE 466 AA; 53192 MW; E4113A5062F58D6E CRC64;

Title: US-09-743-107B-90  
 Perfect score: 66  
 Sequence: 1 CFQWQRXKVR 12

Scoring table: BLOSUM62  
 Gapop 10.0, Gapext 0.5

Run on: February 21, 2003, 07:37:21 / Search time 28.093 Seconds  
 (without alignments)  
 56.918 Million cell updates/sec

Database: A Geneseq\_101002.\*

1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database:

1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	65	98.5	12	21	AA198074 Human lactoferrin
2	64	97.0	12	21	AA198038 Human lactoferrin
3	64	97.0	12	21	AA198046 Human lactoferrin
4	64	97.0	12	21	AA198047 Human lactoferrin
5	64	97.0	12	21	AA198086 Human lactoferrin
6	64	97.0	12	21	AA198089 Human lactoferrin
7	64	97.0	12	21	AA198090 Human lactoferrin
8	64	97.0	13	21	AA198037 Human lactoferrin
9	64	97.0	13	21	AA198048 Human lactoferrin
10	64	97.0	13	21	AA198049 Human lactoferrin

11	64	97.0	14	21	AA198036 Human lactoferrin
12	64	97.0	14	21	AA198050 Human lactoferrin
13	64	97.0	14	21	AA198051 Human lactoferrin
14	64	97.0	15	17	AA198054 Peptide for anti-u
15	64	97.0	15	21	AA198035 Human lactoferrin
16	64	97.0	15	21	AA198062 Human lactoferrin
17	64	97.0	15	21	AA198063 Human lactoferrin
18	64	97.0	16	21	AA198031 Human lactoferrin
19	64	97.0	16	21	AA198064 Human lactoferrin
20	64	97.0	16	21	AA198065 Human lactoferrin
21	64	97.0	17	21	AA198034 Human lactoferrin
22	64	97.0	17	21	AA198066 Human lactoferrin
23	64	97.0	17	21	AA198067 Human lactoferrin
24	64	97.0	18	15	AA198352 Advanced glycosyla
25	64	97.0	18	17	AA198397 Human lactoferrin
26	64	97.0	18	21	AA198033 Human lactoferrin
27	64	97.0	19	21	AA198867 Amino acid sequenc
28	64	97.0	19	21	AA198032 Human lactoferrin
29	64	97.0	20	13	AA198180 Anti microbial pep
30	64	97.0	20	14	AA198441 Lactoferrin-relate
31	64	97.0	20	15	AA198530 Lactoferrin derive
32	64	97.0	20	15	AA198531 Lactoferrin derive
33	64	97.0	20	15	AA198746 Lactoferrin derive
34	64	97.0	20	15	AA198746 Lactoferrin derive
35	64	97.0	20	16	AA198469 Bovine lactoferrin
36	64	97.0	20	16	AA198459 Bovine lactoferrin
37	64	97.0	20	16	AA1980263 Anti-parasitic lac
38	64	97.0	20	16	AA1980264 Anti-parasitic lac
39	64	97.0	20	17	AA198553 Peptide for anti-u
40	64	97.0	20	17	AA198182 Lactoferrin-derive
41	64	97.0	20	17	AA198034 Lactoferrin-derive
42	64	97.0	20	17	AA198067 Lactoferrin-derive
43	64	97.0	20	17	AA198721 Lactoferrin-derive
44	64	97.0	20	17	AA198722 Lactoferrin-derive
45	64	97.0	20	18	AA19826150 Lactoferrin deriva

#### ALIGNMENTS

RESULT 1

AA198074

ID AA198074 standard; Peptide; 12 AA.

AC AA198074;

DT 25-APR-2000 (first entry)

DE Human lactoferrin derived peptide SEQ ID NO:74.

Human; lactoferrin; modification; infection; inflammation; tumour;  
 food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 urinary tract infection; colitis; Candida infection; fungicidal;  
 bactericidal; preservative.

OS Homo sapiens.

OS Synthetic.

XX WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-S601230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.



C fungicidal and bactericidal and may also be used as preservatives.  
C Even though native human lactoferrin have been shown to have desired  
C anti-inflammatory anti-infectious and anti-tumoural properties they  
C cannot be used clinically on a broad basis because of high production  
C costs. Therefore, provision of peptides based on lactoferrin would  
C enable them to be used for the same purposes as lactoferrin at lower  
C cost.

XX Sequence 12 AA;

Query Match 97.0%; Score 64; DB 21; Length 12;  
Best Local Similarity 91.7%; Pred. No. 7.9e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Y 1 CFQWQRMKVR 12  
b 1 CFQWQRMKVR 12

RESULT 4

AY78047  
D AAY78047 standard; Peptide; 12 AA.

X X AAY78047;

X 25-APR-2000 (first entry)

X Human lactoferrin derived peptide SEQ ID NO:47.

X Human; lactoferrin; modification; infection; inflammation; tumour;  
X food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
X urinary tract infection; colitis; Candida infection; fungicidal;  
X bactericidal; preservative.

X Homo sapiens.  
X Synthetic.

X WO200001730-A1.

X 13-JAN-2000.

X 06-JUL-1999; 99WO-SB01230.

X 06-JUL-1998; 98SE-0002441.

X 17-JUL-1998; 98SE-0002562.

X 29-DEC-1998; 98SE-0004614.

X (ASCI-) A+ SCI INVEST AB.

X Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

X WPI; 2000-147388/13.

X New peptides used for treatment and prevention of infections,  
X inflammations and tumors and for use in infant formula food -

X Claim 18; Page 73; 102pp; English.

X AAY78001 to AAY78100 represent peptides having sequences based on human  
X lactoferrin. The peptides are taken up in the intestine through  
X binding to specific lactoferrin receptors and are then transported  
X through the circulation. A medicinal product of the peptide or fragment  
X can be used for treating and/or prevention of infections (such as  
X urinary tract infections, colitis, and Candida infection on a mucosal  
X membrane), inflammations and/or tumors. The peptides can also be used  
X in food stuffs such as infant formula food. The peptides are also  
X fungicidal and bactericidal and may also be used as preservatives.  
X Even though native human lactoferrin have been shown to have desired  
X anti-inflammatory anti-infectious and anti-tumoural properties they  
X cannot be used clinically on a broad basis because of high production  
X costs. Therefore, provision of peptides based on lactoferrin would  
X enable them to be used for the same purposes as lactoferrin at lower  
X cost.

XX Sequence 12 AA;

Query Match 97.0%; Score 64; DB 21; Length 12;  
Best Local Similarity 91.7%; Pred. No. 7.9e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12  
Db 1 CFQWQRMKVR 12

RESULT 5

AY78086  
ID AAY78086 standard; Peptide; 12 AA.

XX AC AAY78086;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:86.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
XX urinary tract infection; colitis; Candida infection; fungicidal;  
XX bactericidal; preservative.

XX Homo sapiens.  
XX Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SB01230.

XX 06-JUL-1998; 98SE-0002441.

XX 17-JUL-1998; 98SE-0002562.

XX 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
XX inflammations and tumors and for use in infant formula food -

XX Claim 22; Page 36; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
XX lactoferrin. The peptides are taken up in the intestine through  
XX binding to specific lactoferrin receptors and are then transported  
XX through the circulation. A medicinal product of the peptide or fragment  
XX can be used for treating and/or prevention of infections (such as  
XX urinary tract infections, colitis, and Candida infection on a mucosal  
XX membrane), inflammations and/or tumors. The peptides can also be used  
XX in food stuffs such as infant formula food. The peptides are also  
XX fungicidal and bactericidal and may also be used as preservatives.  
XX Even though native human lactoferrin have been shown to have desired  
XX anti-inflammatory anti-infectious and anti-tumoural properties they  
XX cannot be used clinically on a broad basis because of high production  
XX costs. Therefore, provision of peptides based on lactoferrin would  
XX enable them to be used for the same purposes as lactoferrin at lower  
XX cost.

XX Sequence 12 AA;

Query Match 97.0%; Score 64; DB 21; Length 12;  
Best Local Similarity 91.7%; Pred. No. 7.9e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQXMRKVR 12  
 DB 1 CFQWQXMRKVR 12

## RESULT 6

AAV78089  
 ID AAY78089 standard; Peptide; 12 AA.

XX AAY78089;  
 AC AAY78089;

XX 25-APR-2000 (first entry)  
 DT 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:89.  
 DE Human lactoferrin; modification; infection; inflammation; tumour;

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.  
 OS Synthetic.

XX WO200001730-A1.  
 PN WO200001730-A1.

XX 13-JAN-2000.  
 PD 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.  
 PF 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.  
 PR 06-JUL-1998; 98SE-0002441.

XX 17-JUL-1998; 98SE-0002562.  
 PR 17-JUL-1998; 98SE-0002562.

XX 29-DEC-1998; 98SE-0004614.  
 PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.  
 PA (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.  
 DR WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 22; Page 37; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 12 AA;  
 SQ Sequence 12 AA;

Query Match 97.0%; Score 64; DB 21; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 7.9e-05;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQXMRKVR 12  
 DB 1 CFQWQXMRKVR 12

## RESULT 7

AAV78090

ID AAY78090 standard; Peptide; 12 AA.

XX AAY78090;  
 AC AAY78090;

XX 25-APR-2000 (first entry)  
 DT 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:90.  
 DE Human lactoferrin; modification; infection; inflammation; tumour;

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.  
 OS Synthetic.

XX WO200001730-A1.  
 PN WO200001730-A1.

XX 13-JAN-2000.  
 PD 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.  
 PF 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.  
 PR 06-JUL-1998; 98SE-0002441.

XX 17-JUL-1998; 98SE-0002562.  
 PR 17-JUL-1998; 98SE-0002562.

XX 29-DEC-1998; 98SE-0004614.  
 PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.  
 PA (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.  
 DR WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 22; Page 37; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 12 AA;  
 SQ Sequence 12 AA;

Query Match 97.0%; Score 64; DB 21; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 7.9e-05;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQXMRKVR 12  
 DB 1 CFQWQXMRKVR 12

## RESULT 8

AAV78037

ID AAY78037 standard; Peptide; 13 AA.

XX AAY78037;  
 AC AAY78037;

XX 25-APR-2000 (first entry)  
 DT 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:37.  
 DE Human lactoferrin derived peptide SEQ ID NO:37.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 PN WO200001730-A1.  
 XX 13-JAN-2000.  
 XX  
 XX 06-JUL-1999; 99WO-SE01230.  
 XX  
 PR 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 PA (ASCII-) A+ SCI INVEST AB.  
 XX  
 PI Hanson LA, Matteby-Baltzer I, Baltzer L, Dolphin GT;  
 XX  
 DR WPI; 2000-147388/13.  
 XX  
 PT New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX  
 PS Claim 12; Page 70; 102pp; English.  
 XX  
 CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 SQ Sequence 13 AA;  
 Query Match 97.0%; Score 64; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 8.6e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 2Y 1 CFQWQRXMKVR 12  
 Db 2 CFQWQRNMRKVR 13  
 RESULT 9  
 AAY78048  
 ID AAY78048 standard; Peptide; 13 AA.  
 AC AAY78048;  
 XX  
 DT 25-APR-2000 (first entry)  
 XX  
 DE Human lactoferrin derived peptide SEQ ID NO:48.  
 XX  
 KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.

OS Synthetic.  
 XX WO200001730-A1.  
 XX 13-JAN-2000.  
 XX  
 XX 06-JUL-1999; 99WO-SE01230.  
 XX  
 PR 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 PA (ASCII-) A+ SCI INVEST AB.  
 XX  
 PI Hanson LA, Matteby-Baltzer I, Baltzer L, Dolphin GT;  
 XX  
 DR WPI; 2000-147388/13.  
 XX  
 PT New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX  
 PS Claim 15; Page 74; 102pp; English.  
 XX  
 CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 SQ Sequence 13 AA;  
 Query Match 97.0%; Score 64; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 8.6e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQWQRXMKVR 12  
 Db 2 CFQWQRNMRKVR 13  
 RESULT 10  
 AAY78049  
 ID AAY78049 standard; Peptide; 13 AA.  
 AC AAY78049;  
 XX  
 DT 25-APR-2000 (first entry)  
 XX  
 DE Human lactoferrin derived peptide SEQ ID NO:49.  
 XX  
 KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200001730-A1.  
 XX  
 PD 13-JAN-2000.  
 XX  
 PF 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 18; Page 74; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 13 AA;

Query Match 97.0%; Score 64; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 8.6e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12  
 DB 2 CFQWQRMKVR 13

RESULT 11

AAAY78036  
 ID AAY78036 standard; Peptide; 14 AA.

XX AC AAY78036;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:36.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.

XX Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

XX 17-JUL-1998; 98SE-0002562.

XX 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX

PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

XX Claim 12; Page 69; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 14 AA;

Query Match 97.0%; Score 64; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 9.2e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12  
 DB 3 CFQWQRMKVR 14

RESULT 12

AAAY78050  
 ID AAY78050 standard; Peptide; 14 AA.

XX AC AAY78050;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:50.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.

XX Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

XX 17-JUL-1998; 98SE-0002562.

XX 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

XX

PS Claim 15; Page 75; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infections and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 14 AA;

Query Match 97.0%; Score 64; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 9.2e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Y 1 CFQWQRMKVR 12  
 |||||  
 Db 3 CFQWQRMKVR 14

RESULT 13

AY78051  
 D AAY78051 standard; Peptide; 14 AA.

AY78051;

25-APR-2000 (first entry)

Human lactoferrin derived peptide SEQ ID NO:51.

Human; lactoferrin; modification; infection; inflammation; tumour;  
 food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 urinary tract infection; colitis; Candida infection; fungicidal;  
 bactericidal; preservative.

Homo sapiens.  
 Synthetic.

WO200001730-A1.

13-JAN-2000.

06-JUL-1999; 99WO-S501230.

06-JUL-1998; 98SE-0002441.

17-JUL-1998; 98SE-0002562.

29-DEC-1998; 98SE-0004614.

(ASCII-) A+ SCI INVEST AB.

Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

WPI; 2000-147388/13.

New peptides used for treatment and prevention of infections,  
 inflammations and tumors and for use in infant formula food

Claim 18; Page 75; 102pp; English.

AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as

CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infections and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 14 AA;

Query Match 97.0%; Score 64; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 9.2e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Y 1 CFQWQRMKVR 12  
 |||||  
 Db 3 CFQWQRMKVR 14

RESULT 14

AY78554  
 ID AAR98554 standard; Peptide; 15 AA.

AY78554;

12-NOV-1996 (first entry)

Peptide for anti-ulcer agent.

anti-ulcer agent; low toxicity; stable; heat-resistant.

Synthetic.

JP08143468-A.

04-JUN-1996.

17-NOV-1994; 94JP-0283869.

17-NOV-1994; 94JP-0283869.

(MORG) MORINAGA MILK IND CO LTD.

WPI; 1996-318857/32.

Anti-ulcer agent contg. peptide - has low toxicity, is  
 heat-resistant and water-soluble

Claim 1; Page 11; 11pp; Japanese.

AAR98531-54 are peptides used in an anti-ulcer agent. The agent is low  
 CC in toxicity, is heat-resistant and stable in aqueous soln.. It can be  
 CC administered orally and be produced in large amounts.

XX Sequence 15 AA;

Query Match 97.0%; Score 64; DB 17; Length 15;  
 Best Local Similarity 91.7%; Pred. No. 9.8e-05;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Y 1 CFQWQRMKVR 12  
 |||||  
 Db 2 CFQWQRMKVR 13

RESULT 15

AY78035  
 ID AAY78035 standard; Peptide; 15 AA.

AY78035;



XX 25-APR-2000 (first entry)  
DT Human lactoferrin derived peptide SEQ ID NO:35.  
XX  
DE Human; lactoferrin; modification; infection; inflammation; tumour;  
XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX  
XX Homo sapiens.  
OS Synthetic.  
OS  
XX WO200001730-A1.  
PN  
XX 13-JAN-2000.  
XX  
XX 06-JUL-1999; 99WO-SE01230.  
PF  
XX 06-JUL-1998; 98SE-0002441.  
PR  
XX 17-JUL-1998; 98SE-0002562.  
PR  
XX 29-DEC-1998; 98SE-0004614.  
XX  
XX (ASCI-) A+ SCI INVEST AB.  
PA  
XX  
XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
PI  
XX WPI; 2000-147388/13.  
DR  
XX  
XX New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
PT  
XX  
XX Claim 12; Page 69; 102pp; English.  
PS  
XX  
XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.  
XX  
XX  
SQ Sequence 15 AA;  
Query Match 97.0%; Score 64; DB 21; Length 15;  
Best Local Similarity 91.7%; Pred.No. 9.8e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CFQWQRXNRKVR 12  
Db 4 CFQWQRXNRKVR 15

Search completed: February 21, 2003, 07:56:43  
Job time : 28.093 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

DM protein - protein search, using sw model

Run on: February 21, 2003, 07:50:40 ; Search time 8.93023 Seconds  
(without alignments)  
39.537 Million cell updates/sec

Title: US-09-743-107B-90

Perfect score: 66

Sequence: 1 CFQWQRXMRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	64	97.0	18	1	US-08-204-487-3
2	64	97.0	18	2	US-08-485-948-8
3	64	97.0	18	2	US-08-628-380-8
4	64	97.0	18	2	US-08-475-055-8
5	64	97.0	20	1	US-07-755-161A-3
6	64	97.0	20	1	US-07-891-174-3
7	64	97.0	20	1	US-08-204-487-1
8	64	97.0	20	1	US-08-256-771-24
9	64	97.0	20	1	US-08-256-771-25
10	64	97.0	20	1	US-08-381-984-24
11	64	97.0	20	1	US-08-381-984-25
12	64	97.0	22	4	US-09-508-734-4
13	64	97.0	22	4	US-09-508-734-6
14	64	97.0	25	1	US-07-755-161A-10
15	64	97.0	25	1	US-07-891-174-10
16	64	97.0	25	1	US-08-204-487-7
17	64	97.0	29	4	US-09-508-734-8
18	64	97.0	36	1	US-07-755-161A-8
19	64	97.0	36	1	US-07-891-174-8
20	64	97.0	36	1	US-08-256-771-30
21	64	97.0	36	1	US-08-381-984-29
22	64	97.0	47	2	US-08-464-182A-6
23	64	97.0	47	2	US-08-406-271-6
24	64	97.0	50	2	US-08-693-274A-7
25	64	97.0	52	4	US-09-017-043A-3
26	64	97.0	53	2	US-08-464-182A-5
27	64	97.0	53	2	US-08-406-271-5

28	64	97.0	54	2	US-08-464-182A-2	Sequence 2, Appli
29	64	97.0	54	2	US-08-406-271-2	Sequence 2, Appli
30	64	97.0	694	3	US-08-724-586-2	Sequence 2, Appli
31	64	97.0	694	4	US-09-421-632-2	Sequence 2, Appli
32	64	97.0	694	4	US-09-932-190-2	Sequence 2, Appli
33	64	97.0	705	2	US-08-655-640-2	Sequence 2, Appli
34	64	97.0	708	2	US-08-655-640-4	Sequence 4, Appli
35	64	97.0	711	1	US-08-354-019-4	Sequence 4, Appli
36	64	97.0	711	1	US-08-461-333-4	Sequence 4, Appli
37	64	97.0	711	3	US-08-464-167-4	Sequence 4, Appli
38	64	97.0	711	3	US-09-158-313-4	Sequence 4, Appli
39	64	97.0	711	4	US-08-476-798-4	Sequence 4, Appli
40	61	92.4	711	1	US-08-145-681-2	Sequence 2, Appli
41	61	92.4	711	1	US-08-250-308-2	Sequence 2, Appli
42	61	92.4	711	1	US-08-453-703-2	Sequence 2, Appli
43	61	92.4	711	2	US-08-456-106-2	Sequence 2, Appli
44	61	92.4	711	3	US-08-456-108-2	Sequence 2, Appli
45	61	92.4	711	4	US-09-365-577-2	Sequence 2, Appli

#### ALIGNMENTS

RESULT 1  
US-08-204-487-3  
; Sequence 3, Application US/08204487  
; Patent No. 5565425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAOKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGEKI  
; APPLICANT: DOSAKO, SHUN'ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION INHIBITORS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESS: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FJN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Peptide  
; FEATURES:  
; NAME/KEY: Peptide  
; LOCATION: 1..18  
; OTHER INFORMATION: /note= "HUMAN LACTOFERRIN PEPTIDE"

OTHER INFORMATION: (20-37)"  
US-08-204-487-3

Query Match 97.0%; Score 64; DB 1; Length 18;  
Best Local Similarity 91.7%; Pred. No. 3.8e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWQRXMRKVR 12  
Db 1 CFQWQRXMRKVR 12

## RESULT 2

US-08-485-948-8  
; Sequence 8, Application US/08485948  
; Patent No. 5855882  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,948  
; FILING DATE:  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/488,217  
; FILING DATE: JUNE 7, 1995  
; APPLICATION NUMBER: 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 947-1-008A  
; REFERENCE/DOCKET NUMBER: 947-1-008A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LF-C1, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; US-08-485-948-8

Query Match 97.0%; Score 64; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 3.8e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWQRXMRKVR 12  
Db 1 CFQWQRXMRKVR 12

## RESULT 3

US-08-628-380-8  
; Sequence 8, Application US/08628380  
; Patent No. 5891341  
; GENERAL INFORMATION:  
; APPLICANT: LI, YONG MING  
; APPLICANT: VLASSARA, HELEN  
; APPLICANT: CERAMI, ANTHONY  
; TITLE OF INVENTION: AGE-MEDIATED INHIBITION OF ANTIBACTERIAL PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/628,380  
; FILING DATE: April 4, 1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LF-C1, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
; US-08-628-380-8

Query Match 97.0%; Score 64; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 3.8e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWQRXMRKVR 12  
Db 1 CFQWQRXMRKVR 12

## RESULT 4

US-08-475-055-8  
; Sequence 8, Application US/08475055  
; Patent No. 5962245  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,055
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/485,948
FILING DATE:
APPLICATION NUMBER: 08/488,217
FILING DATE: JUNE 7, 1995
APPLICATION NUMBER: 08/418,642
FILING DATE: APRIL 7, 1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 947-1-008A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
DESCRIPTION: LF-C1, 8-25
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
US-08-475-055-8

Query Match 97.0%; Score 64; DB 2; Length 18;
Best Local Similarity 91.7%; Pred. No. 3,8e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2Y 1 CFQWQRNMRKVR 12
DB 1 CFQWQRNMRKVR 12

RESULT 5
US-07-755-161A-3
Sequence 3, Application US/07755161A
Patent No. 5304633
GENERAL INFORMATION:
APPLICANT: Mamoru TOMITA et al.
TITLE OF INVENTION: Antimicrobial Peptides and an
TITLE OF INVENTION: Antimicrobial Agent
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500KB
COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/755,161A
FILING DATE: 19910905
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX: 202-371-8856
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY: modified site
LOCATION: 2
IDENTIFICATION METHOD:
OTHER INFORMATION: /note= "thiol group of
OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with
OTHER INFORMATION: thiol group of Cys residue at location 19"
FEATURE:
NAME/KEY: modified site
LOCATION: 19
IDENTIFICATION METHOD:
OTHER INFORMATION: /note= "thiol group of
OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with
OTHER INFORMATION: thiol group of Cys residue at location 2"
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
* RELEVANT RESIDUES IN SEQ ID NO:
US-07-755-161A-3
Query Match 97.0%; Score 64; DB 1; Length 20;
```

Best Local Similarity 91.7%; Pred. No. 4.2e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWQRMKVR 12  
Db 2 CFQWQRMKVR 13

RESULT 6  
US-07-891-174-3  
Sequence 3, Application US/07891174  
Patent No. 5317084  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: Antimicrobial Peptides and an  
TITLE OF INVENTION: Antimicrobial Agent  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/891,174  
FILING DATE: 29-MAY-1992  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/755,161  
FILING DATE: 05-SEP-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX: 202-371-8856  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY: modified site

LOCATION: 2  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 19"  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 19  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 2"  
PUBLICATION INFORMATION:  
AUTHORS:  
JOURNAL:  
TITLE:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-891-174-3

Query Match 97.0%; Score 64; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4.2e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWQRMKVR 12  
Db 2 CFQWQRMKVR 13

RESULT 7  
US-08-204-487-1  
Sequence 1, Application US/08204487  
Patent No. 5585425  
GENERAL INFORMATION:  
APPLICANT: YAWAMOTO, NAOKI  
APPLICANT: NAKASHIMA, HIDEKI  
APPLICANT: MOSUCHI, WATARU  
APPLICANT: TANAKA, SHIGEAKI  
APPLICANT: DOSAKO, SHUN'ICHI  
APPLICANT: KAWASAKI, YOSHITIRO  
APPLICANT: UCHIDA, TOSHIKI  
TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
TITLE OF INVENTION: INHIBITORS  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSER: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
ADDRESSER: THIBEAULT  
STREET: 53 STATE STREET  
CITY: BOSTON  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/204,487  
FILING DATE: 02-MAR-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: CAMPBELL, PAULA A.  
REGISTRATION NUMBER: 32,503  
REFERENCE/DOCKET NUMBER: FJN-019  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..20  
OTHER INFORMATION: /note= "ANTIBACTERIAL PEPTIDE  
DERIVED FROM HUMAN LACTOPERRIN"  
US-08-204-487-1

Query Match 97.0%; Score 64; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4.2e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWRXMRKVR 12  
b 2 CFQWRXMRKVR 13

RESULT 8  
US-08-256-771-24  
Sequence 24, Application US/08256771  
Patent No. 5658591  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
TITLE OF INVENTION: PRODUCTS THEREWITH  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,771  
FILING DATE: July 22, 1994  
CLASSIFICATION: 514  
PRIOR APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "Cys residues are linked by  
OTHER INFORMATION: disulfide bond"

US-08-256-771-24  
Query Match 97.0%; Score 64; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4.2e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWRXMRKVR 12  
Db 2 CFQWRXMRKVR 13

RESULT 9  
US-08-256-771-25  
Sequence 25, Application US/08256771  
Patent No. 5658591  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
TITLE OF INVENTION: PRODUCTS THEREWITH  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,771  
FILING DATE: July 22, 1994  
CLASSIFICATION: 514  
PRIOR APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "Cys residues are protected to  
OTHER INFORMATION: prevent disulfide bond"  
US-08-256-771-25

Query Match 97.0%; Score 64; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4.2e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWRXMRKVR 12  
Db 2 CFQWRXMRKVR 13

RESULT 10  
US-08-381-984-24

Sequence 24, Application US/08381984  
Patent No. 5804555  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIOXIDANT  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/381,984  
FILING DATE: April 11, 1995  
CLASSIFICATION: 252  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:

## INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "cysteine residues at positions 2  
OTHER INFORMATION: and 19 are bonded by disulfide linkage"

## FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "the specified peptide as well as

OTHER INFORMATION: peptides including the specified peptide as a fragment thereof

US-08-381-984-24

Query Match 97.0%; Score 64; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4.2e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12

DB 2 CFQWQRMKVR 13

## RESULT 11

US-08-381-984-25  
Sequence 25, Application US/08381984  
Patent No. 5804555  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIOXIDANT  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack

STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/381,984  
FILING DATE: April 11, 1995  
CLASSIFICATION: 252  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:

## INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "the specified peptide as well as

OTHER INFORMATION: peptides including the specified peptide as a fragment thereof

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "cysteine residues at positions 2

OTHER INFORMATION: and 19 are chemically modified to prevent disulfide linkage"

US-08-381-984-25

Query Match 97.0%; Score 64; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 4.2e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12

DB 2 CFQWQRMKVR 13

## RESULT 12

US-09-508-734-4  
Sequence 4, Application US/09508734  
Patent No. 6423509  
GENERAL INFORMATION:  
APPLICANT: Sanyang Genex Corporation  
TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and  
FILE OF INVENTION: useful microorganism thereof  
FILE REFERENCE: PA/SYG/00139  
CURRENT APPLICATION NUMBER: US/09/508,734  
CURRENT FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: PCT/KE99/00373  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: KR1998-29351  
PRIOR FILING DATE: 1998-07-13  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Kopatentin 1.71  
SEQ ID NO 4

```

; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
JS-09-508-734-4
Query Match          97.0%; Score 64; DB 4; Length 22;
Best Local Similarity 91.7%; Pred. No. 4.6e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2Y 1 CFQWQRMKRVK 12
   |||||
2b 2 CFQWQRMKRVK 13
   |||||

RESULT 13
JS-09-508-734-6
; Sequence 6, Application US/09508734
; Patent No. 6423509
; GENERAL INFORMATION:
; APPLICANT: Samyang Genex Corporation
; TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and
; FILE REFERENCE: PA/SYG/00139
; CURRENT APPLICATION NUMBER: US/09/508,734
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/KR99/00373
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: KR1998-29351
; PRIOR FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Kopatentin 1.71
; SEQ ID NO 6
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
JS-09-508-734-6

Query Match          97.0%; Score 64; DB 4; Length 24;
Best Local Similarity 91.7%; Pred. No. 5e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2Y 1 CFQWQRMKRVK 12
   |||||
2b 3 CFQWQRMKRVK 14
   |||||

RESULT 14
JS-07-755-161A-10
; Sequence 10, Application US/07755161A
; Patent No. 5304633
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; TITLE OF INVENTION: Antimicrobial Agent
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/755,161A
; FILING DATE: 19910905
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 4
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 21"
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 21
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 4"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-07-755-161A-10

Query Match          97.0%; Score 64; DB 1; Length 25;
Best Local Similarity 91.7%; Pred. No. 5.2e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMKRVK 12
   |||||
Db 4 CFQWQRMKRVK 15
   |||||
```



```
RESULT 15
US-07-891-174-10
: Sequence 10, Application US/07891174
: Patent No. 5317084
: GENERAL INFORMATION:
: APPLICANT: Mamoru TOMITA et al.
: TITLE OF INVENTION: Antimicrobial Peptides and an
: TITLE OF INVENTION: Antimicrobial Agent
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Wenderoth, Lind & Ponack
: STREET: 805 Fifteenth Street, N.W., #700
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: MS-DOS
: SOFTWARE: DisplayWrite
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/891,174
: FILING DATE: 29-MAY-1992
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/755,161
: FILING DATE: 05-SEP-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Warren M. Cheek Jr.
: REGISTRATION NUMBER: 33,367
: REFERENCE/DOCKET NUMBER:
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-371-8850
: TELEFAX: 202-371-8856
: TELEX:
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 25 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE:
: HYPOTHETICAL:
: ANTI-SENSE:
: FRAGMENT TYPE:
: ORIGINAL SOURCE:
: ORGANISM:
: STRAIN:
: INDIVIDUAL ISOLATE:
: DEVELOPMENTAL STAGE:
: HAPLOTYPE:
: TISSUE TYPE:
: CELL TYPE:
: CELL LINE:
: ORGANELLE:
: IMMEDIATE SOURCE:
: LIBRARY:
: CLONE:
: POSITION IN GENOME:
: CHROMOSOME/SEGMENT:
: MAP POSITION:
: UNITS:
: FEATURE:
: NAME/KEY: modified site
: LOCATION: 4
: IDENTIFICATION METHOD:
: OTHER INFORMATION: /note= "thiol group of
: OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with
: OTHER INFORMATION: thiol group of Cys residue at location 21"
: FEATURE:
: NAME/KEY: modified site
: LOCATION: 21

: IDENTIFICATION METHOD:
: OTHER INFORMATION: /note= "thiol group of
: OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with
: OTHER INFORMATION: thiol group of Cys residue at location 4"
: PUBLICATION INFORMATION:
: AUTHORS:
: TITLE:
: JOURNAL:
: VOLUME:
: ISSUE:
: PAGES:
: DATE:
: DOCUMENT NUMBER:
: FILING DATE:
: PUBLICATION DATE:
: RELEVANT RESIDUES IN SEQ ID NO:
: US-07-891-174-10
: Query Match 97.0%; Score 64; DB 1; Length 25;
: Best Local Similarity 91.7%; Pred. No. 5.2e-05;
: Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
: QY 1 CFQWQRXMRKVR 12
: ||||| |||||
: Db 4 CFQWQRXMRKVR 15
: Search completed: February 21, 2003, 08:04:26
: Job time : 8.93023 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 08:08:15 ; Search time 6.88372 Seconds  
(without alignments)  
54.162 Million cell updates/sec

Title: US-09-743-107b-90

Perfect score: 66

Sequence: 1 CFQWRXMRKVR 12

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 156504 seqs, 31069816 residues

Total number of hits satisfying chosen parameters: 156504

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	64	97.0	15	9	US-09-798-869-2
2	64	97.0	25	9	US-09-798-869-20
3	64	97.0	694	9	US-10-023-096-2
4	56	84.8	15	9	US-09-798-869-6
5	53	80.3	15	9	US-09-798-869-3
6	53	80.3	25	9	US-09-798-869-23
7	45	68.2	15	9	US-09-798-869-7
8	44	66.7	15	9	US-09-798-869-4
9	44	66.7	25	9	US-09-798-869-22
10	41	62.1	15	9	US-09-798-869-8
11	41	62.1	15	9	US-09-798-869-29
12	41	62.1	15	9	US-09-798-869-30
13	37	56.1	21	10	US-09-864-761-47985
14	37	56.1	489	9	US-09-888-320-2
15	36	54.5	209	10	US-09-904-536-8
16	36	54.5	209	10	US-09-904-536-9
17	36	54.5	209	10	US-09-904-536-11
18	36	54.5	209	10	US-09-904-536-12
19	36	54.5	209	10	US-09-904-536-13

20	36	54.5	209	10	US-09-904-536-14	Sequence 14, Appl
21	36	54.5	209	10	US-09-904-536-15	Sequence 15, Appl
22	36	54.5	209	10	US-09-904-536-16	Sequence 16, Appl
23	36	54.5	209	10	US-09-904-536-17	Sequence 17, Appl
24	36	54.5	209	10	US-09-904-536-18	Sequence 18, Appl
25	36	54.5	212	10	US-09-904-536-10	Sequence 10, Appl
26	36	54.5	235	9	US-10-095-449-6	Sequence 6, Appl
27	36	54.5	235	10	US-09-448-378-1	Sequence 1, Appl
28	36	54.5	235	10	US-09-983-806-6	Sequence 6, Appl
29	36	54.5	235	10	US-09-904-536-1	Sequence 1, Appl
30	35	53.0	86	9	US-09-738-626-5715	Sequence 5715, Ap
31	34	51.5	95	10	US-09-764-864-1031	Sequence 1031, Ap
32	34	51.5	846	9	US-10-051-409-4	Sequence 4, Appl
33	34	51.5	1617	9	US-10-050-453A-2	Sequence 2, Appl
34	33	50.0	15	9	US-09-798-869-5	Sequence 5, Appl
35	33	50.0	46	10	US-09-864-761-48879	Sequence 48879, A
36	33	50.0	579	10	US-09-824-588-2	Sequence 2, Appl
37	32.5	49.2	124	10	US-09-864-761-44327	Sequence 44327, A
38	32	48.5	13	9	US-09-798-869-26	Sequence 26, Appl
39	32	48.5	14	9	US-09-798-869-25	Sequence 25, Appl
40	32	48.5	15	9	US-09-798-869-1	Sequence 1, Appl
41	32	48.5	15	9	US-09-798-869-10	Sequence 10, Appl
42	32	48.5	15	9	US-09-798-869-28	Sequence 28, Appl
43	32	48.5	18	9	US-09-798-869-24	Sequence 24, Appl
44	32	48.5	25	9	US-09-798-869-21	Sequence 21, Appl
45	32	48.5	25	10	US-09-903-652-4	Sequence 4, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-798-869-2  
; Sequence 2, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-09-798-869-2

Query Match 97.0%; Score 64; DB 9; Length 15;  
Best Local Similarity 91.7%; Pred. No. 2.5e-05;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWRXMRKVR 12

DB 3 CFQWRXMRKVR 14

##### RESULT 2

US-09-798-869-20  
; Sequence 20, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON

```

Query Match          97.0%;   Score 64;   DB 9;   Length 694;
Best Local Similarity 91.7%;   Pred. No. 0.00096;
Matches 11;   Conservative 0;   Mismatches 1;   Indels 0;   Gaps 0;

QY      1 CFQWQRXMRKVR 12
Db      22 CFQWQRNMRKVR 33

RESULT 4
US-09-798-869-6
; Sequence 5, Application US/09798869
; Publication No. US20030022821A1
; GENERAL INFORMATION:
; APPLICANT: JOHN SIGURD SVENDSEN
; APPLICANT: (YSTEIN REKDAL
; APPLICANT: BALDUR SVEINBJ(RNSSON
; APPLICANT: LARS VORLAND
; TITLE OF INVENTION: BIOACTIVE PEPTIDES
; FILE REFERENCE: A34049-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/798,869
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: PCT/GB99/02851
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: GB9818938.4
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of Homo sapiens
; OTHER INFORMATION: sequence)
US-09-798-869-6

Query Match          84.8%;   Score 56;   DB 9;   Length 15;
Best Local Similarity 83.3%;   Pred. No. 0.00062;
Matches 10;   Conservative 0;   Mismatches 2;   Indels 0;   Gaps 0;

QY      1 CFQWQRXMRKVR 12
Db      3 CFQWQRNMRKVR 14

RESULT 5
US-09-798-869-3
; Sequence 3, Application US/09798869
; Publication No. US20030022821A1
; GENERAL INFORMATION:
; APPLICANT: JOHN SIGURD SVENDSEN
; APPLICANT: (YSTEIN REKDAL
; APPLICANT: BALDUR SVEINBJ(RNSSON
; APPLICANT: LARS VORLAND
; TITLE OF INVENTION: BIOACTIVE PEPTIDES
; FILE REFERENCE: A34049-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/798,869
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: PCT/GB99/02851
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: GB9818938.4
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 15
; TYPE: PRT
; ORGANISM: CAPRINE
US-09-798-869-3

Query Match          80.3%;   Score 53;   DB 9;   Length 15;
Best Local Similarity 72.7%;   Pred. No. 0.00021;
Matches 11;   Conservative 0;   Mismatches 1;   Indels 0;   Gaps 0;

```

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Y 1 CFQQRXMKV 11  
|:|||||:  
b 3 CYQQRMRKL 13

RESULT 6  
US-09-798-869-23  
Sequence 23, Application US/09798869  
Publication No. US20030022821A1  
GENERAL INFORMATION:  
APPLICANT: JOHN SIGURD SVENDSEN  
APPLICANT: (YSTEIN REKDAL  
APPLICANT: BALDUR SVEINBJ (RNSSON  
APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 23  
LENGTH: 25  
TYPE: PRT  
ORGANISM: CAPRINE  
US-09-798-869-23

Query Match 80.3%; Score 53; DB 9; Length 25;  
Best Local Similarity 72.7%; Pred. No. 0.0034;  
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Y 1 CFQQRXMKV 11  
|:|||||:  
b 3 CYQQRMRKL 13

RESULT 7  
US-09-798-869-7  
Sequence 7, Application US/09798869  
Publication No. US20030022821A1  
GENERAL INFORMATION:  
APPLICANT: JOHN SIGURD SVENDSEN  
APPLICANT: (YSTEIN REKDAL  
APPLICANT: BALDUR SVEINBJ (RNSSON  
APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 7  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: synthetic peptide (modified form of homo sapiens  
OTHER INFORMATION: sequence)

Query Match 68.2%; Score 45; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.051;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Y 1 CFQQRXMKV 11  
|:|||||:  
b 3 CYQQRMRKL 13

RESULT 8  
US-09-798-869-4  
Sequence 4, Application US/09798869  
Publication No. US20030022821A1  
GENERAL INFORMATION:  
APPLICANT: JOHN SIGURD SVENDSEN  
APPLICANT: (YSTEIN REKDAL  
APPLICANT: BALDUR SVEINBJ (RNSSON  
APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 15  
TYPE: PRT  
ORGANISM: MURINE  
US-09-798-869-4

Query Match 66.7%; Score 44; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.077;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Y 1 CFQQRXMKV 11  
|:|||||:  
b 3 CLRQWNRKV 13

RESULT 9  
US-09-798-869-22  
Sequence 22, Application US/09798869  
Publication No. US20030022821A1  
GENERAL INFORMATION:  
APPLICANT: JOHN SIGURD SVENDSEN  
APPLICANT: (YSTEIN REKDAL  
APPLICANT: BALDUR SVEINBJ (RNSSON  
APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 22  
LENGTH: 25  
TYPE: PRT  
ORGANISM: MURINE  
US-09-798-869-22

Query Match 66.7%; Score 44; DB 9; Length 25;  
Best Local Similarity 63.6%; Pred. No. 0.12;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Y 1 CFQQRXMKV 11  
|:|||||:  
b 3 CLRQWNRKV 13

RESULT 10  
US-09-798-869-8  
; Sequence 8, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON  
; APPLICANT: LARS VORLAND  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of caprine  
; OTHER INFORMATION: sequence)  
US-09-798-869-8

Query Match 62.1%; Score 41; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.26;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQRMKRV 11  
|:|:|:|:|:  
Db 3 CLRWQEMKVL 13

RESULT 11  
US-09-798-869-29  
; Sequence 29, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 29  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: BOVINE  
US-09-798-869-29

Query Match 62.1%; Score 41; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.26;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQRMKRV 11  
|:|:|:|:|:  
Db 3 CLRWQEMKVL 13

RESULT 12  
US-09-798-869-30

; Sequence 30, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 30  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: BOVINE  
US-09-798-869-30

Query Match 62.1%; Score 41; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.26;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQRMKRV 11  
|:|:|:|:|:  
Db 3 CLRWQEMKVL 13

RESULT 13  
US-09-864-761-47985  
; Sequence 47985, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecmca-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: US 60/234,697  
 ; PRIOR FILING DATE: 2000-09-21  
 ; PRIOR APPLICATION NUMBER: US 09/608,408  
 ; PRIOR FILING DATE: 2000-06-30  
 ; PRIOR APPLICATION NUMBER: US 09/774,203  
 ; PRIOR FILING DATE: 2001-01-29  
 ; NUMBER OF SEQ ID NOS: 49117  
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
 ; SEQ ID NO 47985  
 ; LENGTH: 21  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:

OTHER INFORMATION: MAP TO AL096701.14  
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3  
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1  
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6  
 OTHER INFORMATION: EST\_HUMAN HIT: AW294800.1, EVALUATE 1.00e-06  
 IS-09-864-761-47985

Query Match 56.1%; Score 37; DB 10; Length 21;  
 Best Local Similarity 83.3%; Pred. No. 1.8;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Y 1 CFQWOR 6  
 |||||  
 b 16 CFQWR 21

## RESULT 14

IS-09-888-320-2  
 Sequence 2, Application US/09888320  
 Publication No. US20030013090A1  
 GENERAL INFORMATION:  
 APPLICANT: Barry III, Clifton E.  
 APPLICANT: DeBarber, Andrea E.  
 APPLICANT: Mdulili, Khisimuzi  
 APPLICANT: Bekker, Linda-Gail  
 APPLICANT: The Government of the United States of America  
 APPLICANT: as represented by The Secretary of the  
 APPLICANT: Department of Health and Human Services  
 TITLE OF INVENTION: Methods of Diagnosing Multidrug Resistant Tuberculosis  
 FILE REFERENCE: 015280-413100US  
 CURRENT APPLICATION NUMBER: US/09/888,320  
 CURRENT FILING DATE: 2001-06-22  
 PRIOR FILING DATE: 2001-06-22  
 PRIOR APPLICATION NUMBER: US 60/214,187  
 NUMBER OF SEQ ID NOS: 16  
 SOFTWARE: PatentIn ver. 2.1  
 SEQ ID NO 2  
 LENGTH: 489  
 TYPE: PRT  
 ORGANISM: Mycobacterium tuberculosis  
 OTHER INFORMATION: wild-type EtaA monooxygenase (Rv3854c, EthA)  
 IS-09-888-320-2

Query Match 56.1%; Score 37; DB 9; Length 489;  
 Best Local Similarity 54.5%; Pred. No. 35;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Y 1 CFQWORXMRKV 11  
 |||||  
 b 253 CQKWRFRMRKM 263

## RESULT 15

S-09-904-536-8  
 Sequence 8, Application US/09904536  
 Patent No. US2002011475A1  
 GENERAL INFORMATION:

; APPLICANT: Graddis, Thomas J.  
 ; APPLICANT: McGrew, Jeffrey T.  
 ; TITLE OF INVENTION: FLT3-L MUTANTS AND METHODS OF USE  
 ; FILE REFERENCE: 03260.0028  
 ; CURRENT APPLICATION NUMBER: US/09/904,536  
 ; CURRENT FILING DATE: 2001-07-16  
 ; PRIOR APPLICATION NUMBER: PRIOR APPLICATION: 09/109,100  
 ; PRIOR FILING DATE: 1999-07-02  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: PatentIn ver. 2.1  
 ; SEQ ID NO 8  
 ; LENGTH: 209  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-904-536-8

Query Match 54.5%; Score 36; DB 10; Length 209;  
 Best Local Similarity 50.0%; Pred. No. 23;  
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWORXMRK 10  
 |||||  
 Db 178 CLHWQRTTTR 187

Search completed: February 21, 2003, 08:11:56  
 Job time : 7.88372 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:48:01 ; Search time 10.6047 Seconds  
(without alignments)  
108.784 Million cell updates/sec

Title: US-09-743-107b-90

Perfect score: 66

Sequence: 1 CFQWRXMRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	97.0	711	1	TFHUL
2	53	80.3	708	2	lactotransferrin p
3	50	75.8	33	3	lactoferrin - goat
4	44	66.7	707	1	lactoferrin - sheep
5	43	65.2	4568	2	lactoferrin precursor
6	40	60.6	275	2	dynamin beta heavy
7	40	60.6	932	2	hypothetical prote
8	39	59.1	206	2	hypothetical prote
9	39	59.1	206	2	pyridoxamine 5'-ph
10	39	59.1	208	2	pyridoxamine 5'-ph
11	38	57.6	289	2	probable pyridoxam
12	38	57.6	515	2	33.3K hypothetical
13	38	57.6	543	2	probable cytochrom
14	38	57.6	1135	2	cytochrome P450 ho
15	37	56.1	85	2	phytochrome C - so
16	37	56.1	121	2	polymyxin B resist
17	37	56.1	255	2	hypothetical prote
18	37	56.1	433	2	signal peptidase I
19	37	56.1	489	2	hypothetical prote
20	37	56.1	584	2	probable monooxyge
21	36	54.5	235	2	flit3 ligand - huma
22	36	54.5	275	1	interleukin-2 rece
23	36	54.5	275	1	interleukin-2 rece
24	36	54.5	456	2	hypothetical prote
25	36	54.5	456	2	hypothetical prote
26	36	54.5	1391	2	hypothetical prote
27	36	54.5	1397	2	protein F09C3.1 li
28	35	53.0	124	2	F1511.22 (imported
29	35	53.0	220	1	pyridoxamine-phosp

fprA protein - Myx  
hypothetical prote  
cell cycle arrest  
ornithine decarbox  
probable cytochrom  
cytochrome P450-11  
cytochrome P450 ho  
cytochrome P450-11  
cytochrome P450 ho  
hypothetical prote  
hypothetical prote  
finger protein unk  
hypothetical prote  
kinase-related pro  
protein F21H11.2 l  
carcinoembryonic a

ALIGNMENTS

RESULT 1

TFHUL

Lactotransferrin precursor [validated] - human

N:Alternate names: lactoferrin

C:Species: Homo sapiens (man)

C>Date: 31-Mar-1992 #sequence revision 21-Nov-1997 #text change 08-Dec-2000

C:Accession: G01394; S11228; A45401; S10324; S15853; S20841; S07160; A51169; A31000; S74

R:Cho, Y.

Submitted to the EMBL Data Library, March 1994

A:Reference number: G06820

A:Accession: G01394

A>Status: Preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-711 <CHO>

A:Cross-references: EMBL:U07643; NID:g467236; PIDN:AAB60324.1; PID:g467237

R:Rey, M.M.; Woloshuk, S.L.; deBoer, H.A.; Pieper, F.R.

Nucleic Acids Res. 18, 5288, 1990

A>Title: Complete nucleotide sequence of human mammary gland lactoferrin.

A:Reference number: S11228; MUID:90384839; PMID:2402455

A:Accession: S11228

A:Molecule type: mRNA

A:Residues: 1-148; T, 150-422, C, 424-711 <REY>

A:Cross-references: EMBL:X53961; NID:g34415; PIDN:CAA37914.1; PID:g34416

R:Teng, C.T.; Liu, Y.; Yang, N.; Walmer, D.; Panella, T.

Mol. Endocrinol. 6, 1969-1981, 1992

A>Title: Differential molecular mechanism of the estrogen action that regulates lactofer

A:Reference number: A45401; MUID:93125571; PMID:1480183

A:Accession: A45401

A:Molecule type: DNA

A:Residues: 1-15 <TEN>

A:Cross-references: GB:S52659; NID:g263311; PIDN:AAB24877.1; PID:g263312

A:Experimental source: placenta

A>Note: sequence extracted from NCBI backbone (NCBIP:122202)

R:Powell, N.J.; Ogden, J.E.

Nucleic Acids Res. 18, 4013, 1990

A>Title: Nucleotide sequence of human lactoferrin cDNA.

A:Reference number: S10324; MUID:90326549; PMID:2374734

A:Accession: S10324

A:Molecule type: mRNA

A:Residues: 1-711 <POW>

A:Cross-references: EMBL:X52941; NID:g34411; PIDN:CAA37116.1; PID:g34412

R:Stowell, K.W.; Rado, T.A.; Funk, W.D.; Tweedie, J.W.

Biochem. J. 276, 349-355, 1991

A>Title: Expression of cloned human lactoferrin in baby-hamster kidney cells.

A:Reference number: S15853; MUID:91264786; PMID:2049066

A:Accession: S15853

A>Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 20-31 <ST1>

A:Accession: S20841

A:Molecule type: protein

A:Residues: 20-26, X, 30-31 <ST2>

R;Rado, T.A.; Wei, X.; Benz Jr., E.J.  
 Blood 70, 989-993, 1987  
 A;Title: Isolation of lactoferrin cDNA from a human myeloid library and expression of mRNA  
 A;Reference number: S07160; MUID:88001031; PMID:3477300  
 A;Accession: S07160  
 A;Molecule type: mRNA  
 A;Residues: 436-487, 'A', 499-711 <RAD>  
 A;Cross-references: EMBL:M18642; NID:AAA86665.1; PID:G386855  
 R;Panella, T.J.; Liu, Y.; Huang, A.T.; Teng, C.T.  
 Cancer Res. 51, 3037-3043, 1991  
 A;Title: Polymorphism and altered methylation of the lactoferrin gene in normal leukocytes  
 A;Reference number: A61169; MUID:91325214; PMID:1674448  
 A;Accession: A61169  
 A;Status: not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 3-701, 'GWKPNV' <PAN>  
 A;Experimental source: normal breast tissue  
 R;Metz-Boutigue, M.H.; Jolles, J.; Mazurier, J.; Schoentgen, F.; Legrand, D.; Spik, G.;  
 Eur. J. Biochem. 145, 659-666, 1984  
 A;Title: Human lactotransferrin: amino acid sequence and structural comparisons with other  
 A;Reference number: A31000; MUID:85076667; PMID:6510420  
 A;Accession: A31000  
 A;Molecule type: protein  
 A;Residues: 20-140, 142-169, 171-203, 'L', 205, 'K', 207-208, 'K', 210-385, 'Q', 387-391, 'W', 393-4  
 A;Note: this is the final paper in a series  
 R;Houen, G.; Hoegedal, E.V.; Barkholt, V.; Norrskov, L.  
 Eur. J. Biochem. 241, 303-308, 1996  
 A;Title: Lactoferrin: similarity to diamine oxidase and purification by aminohexyl affinity  
 A;Reference number: S74119; MUID:97054624; PMID:8898921  
 A;Accession: S74119  
 A;Molecule type: protein  
 A;Residues: 'Q', 23-24, 'R', 26-27, 'XX', 30-32 <HOU>  
 A;Experimental source: neutrophil granulocytes  
 C;Genetics:  
 A;Gene: GDB:LTF  
 A;Cross-references: GDB:119368; OMIM:150210  
 A;Map position: 3q21-3q23  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein; iron binding; milk  
 F;1-19/Domain: signal sequence #status predicted <SIG>  
 F;20-711/Product: lactotransferrin #status experimental <MAT>  
 F;360-699/Domain: transferrin repeat homology <TRH2>  
 F;29-65, 39-56, 135-218, 177-190, 201, 251-265, 503-697, 595-609/Disulfide bonds: #status e  
 F;157, 498/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F;368-400, 378-391, 425-706, 447-669, 479-554, 513-527, 524-537, 647-652/Disulfide bonds: #stat

Query Match 97.0%; Score 64; DB 1; Length 711;  
 Best Local Similarity 91.7%; Pred. No. 0.00063;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12  
 DB 39 CFQWQRMKVR 50

RESULT 2  
 J02323  
 lactoferrin - goat  
 C;Species: Capra aegagrus hircus (domestic goat)  
 C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 07-May-1999  
 C;Accession: J02323  
 R;Le Provost, F.; Nocart, M.; Guerin, G.; Martin, P.  
 Biochem. Biophys. Res. Commun. 203, 1324-1332, 1994  
 A;Title: Characterization of the goat lactoferrin cDNA: assignment of the relevant locus  
 A;Reference number: J02323; MUID:94380047; PMID:8093048  
 A;Accession: J02323  
 A;Molecule type: mRNA  
 A;Residues: 1-708 <LEP>  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein  
 F;359-696/Domain: transferrin repeat homology <TRH2>  
 F;252, 300, 387, 495, 564/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 80.3%; Score 53; DB 2; Length 708;  
 Best Local Similarity 72.7%; Pred. No. 0.083;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMKRV 11  
 DB 38 CFQWQRMKRL 48

RESULT 3  
 S52107  
 lactoferrin - sheep (fragment)  
 C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C;Date: 14-Jul-1995 #sequence\_revision 05-Dec-1997 #text\_change 05-Dec-1997  
 C;Accession: S52107  
 R;Qian, Z.Y.; Jolles, P.; Migliore-Samour, D.; Fiat, A.M.  
 Biochim. Biophys. Acta 1243, 25-32, 1995  
 A;Title: Isolation and characterization of sheep lactoferrin, an inhibitor of platelet  
 A;Reference number: S52107; MUID:95127729; PMID:7827104  
 A;Accession: S52107  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-33 <QIA>  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication

Query Match 75.8%; Score 50; DB 2; Length 33;  
 Best Local Similarity 63.6%; Pred. No. 0.016;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRMKRV 11  
 DB 19 CFQWQRMKRL 29

RESULT 4  
 A28438  
 lactoferrin precursor - mouse  
 N;Alternate names: lactotransferrin  
 C;Species: Mus musculus (house mouse)  
 C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C;Accession: A28438; A41205  
 R;Pentecost, B.T.; Teng, C.T.  
 J. Biol. Chem. 262, 10134-10139, 1987  
 A;Title: Lactotransferrin is the major estrogen inducible protein of mouse uterine secr  
 A;Reference number: A92596; MUID:87280033; PMID:3611056  
 A;Accession: A28438  
 A;Molecule type: mRNA  
 A;Residues: 3-707 <PEN>  
 A;Cross-references: EMBL:J03298  
 R;Liu, Y.; Teng, C.T.  
 J. Biol. Chem. 266, 21880-21885, 1991  
 A;Title: Characterization of estrogen-responsive mouse lactoferrin promoter.  
 A;Reference number: A41205; MUID:92042099; PMID:1939212  
 A;Accession: A41205  
 A;Molecule type: DNA  
 A;Residues: 1-15 <LIU>  
 A;Cross-references: GB:M74778  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein  
 F;1-19/Domain: signal sequence #status predicted <SIG>  
 F;20-707/Product: lactotransferrin #status predicted <MAT>  
 F;358-695/Domain: transferrin repeat homology <TRH2>  
 F;494/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 66.7%; Score 44; DB 1; Length 707;  
 Best Local Similarity 63.6%; Pred. No. 3.6;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQRMKRV 11  
 DB 37 CLRWQNMKRV 47





A;Cross-references: GB:AE008688; PIDN:AAL41776.1; PID:g17739129; GSPDB:GN00186  
 A;Experimental source: strain C58 (Dupont)  
 C;Genetics:  
 A;Gene: pdxH  
 A;Map position: circular chromosome  
 C;Superfamily: pyridoxamine-phosphate oxidase

Query Match 59.1%; Score 39; DB 2; Length 206;  
 Best Local Similarity 50.0%; Pred. No. 9;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWQXMRKVR 12  
 |||:|:  
 Db 88 CFHWKSLRQVR 99

RESULT 10  
 AG3441  
 probable pyridoxamine-phosphate oxidase (EC 1.4.3.5) [imported] - Brucella melitensis (s  
 C;Species: Brucella melitensis  
 C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 03-May-2002  
 C;Accession: AG3441  
 R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova,  
 ; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess  
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
 A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
 A;Reference number: AD3252; PMID:11756688  
 A;Accession: AG3441  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-208 <KUR>  
 A;Cross-references: GB:AE008917; PIDN:AAL52698.1; PID:g17983525; GSPDB:GN00190  
 A;Experimental source: strain 16M  
 C;Genetics:  
 A;Gene: BWE11517  
 A;Map position: I  
 C;Superfamily: Pyridoxamine-phosphate oxidase  
 C;Keywords: oxidoreductase

Query Match 59.1%; Score 39; DB 2; Length 208;  
 Best Local Similarity 50.0%; Pred. No. 9;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWQXMRKVR 12  
 |||:|:  
 Db 90 CFHWKSLRQVR 101

RESULT 11  
 G86403  
 33.3K hypothetical protein F28L5.13 - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Nov-2001  
 C;Accession: G86403  
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A;Reference number: AB6141; MUID:21016719; PMID:11130712  
 A;Accession: G86403  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-289 <STO>  
 A;Cross-references: GB:AE005172; NID:g10998925; PIDN:AG26065.1; GSPDB:GN00141  
 C;Genetics:  
 A;Map position: 1

Query Match 57.6%; Score 38; DB 2; Length 289;  
 Best Local Similarity 50.0%; Pred. No. 19;  
 Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CFQWQXMRKVR 12  
 |||:|:  
 Db 8 CFTWEYARHVR 19

RESULT 12  
 T00510  
 probable cytochrome P450 At2g23220 [imported] - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 23-Mar-2001  
 C;Accession: T00510; A84622  
 R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kau  
 submitted to the EMBL Data Library, November 1997  
 A;Description: Arabidopsis thaliana chromosome II BAC T20D16 genomic sequence.  
 A;Reference number: Z14164  
 A;Accession: T00510  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-515 <ROU>  
 A;Cross-references: EMBL:AC002391; NID:g2642427; PID:g2642441  
 A;Experimental source: cultivar Columbia  
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,  
 Nature 402, 761-768, 1999  
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A;Reference number: A84420; MUID:20083487; PMID:10617197  
 A;Accession: A84622  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-515 <STO>  
 A;Cross-references: GB:AE002093; NID:g2642441; PIDN:AAB87109.1; GSPDB:GN00139  
 C;Genetics:  
 A;Gene: At2g23220; T20D16.15  
 A;Map position: 2  
 A;Introns: 182/3; 310/3  
 C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
 C;Keywords: heme; iron; metalloprotein  
 F;312-471/Domains: cytochrome P450 homology <P45>  
 F;449/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 57.6%; Score 38; DB 2; Length 515;  
 Best Local Similarity 83.3%; Pred. No. 33;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQ 6  
 |||:|:  
 Db 469 CFQWQ 474

RESULT 13  
 T00513  
 cytochrome P450 homolog At2g23190 - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 23-Mar-2001  
 C;Accession: T00513; F84621  
 R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kau  
 submitted to the EMBL Data Library, November 1997  
 A;Description: Arabidopsis thaliana chromosome II BAC T20D16 genomic sequence.  
 A;Reference number: Z14164  
 A;Accession: T00513  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-543 <ROU>  
 A;Cross-references: EMBL:AC002391; NID:g2642427; PID:g2642444  
 A;Experimental source: cultivar Columbia  
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,

Nature 402, 761-768, 1999  
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A/Reference number: R84420; MUID:20083487; PMID:10617197  
A/Accession: F84621  
A/Status: Preliminary  
A/Molecule type: DNA  
A/Residues: 1-543 <STO>  
A/Cross-references: GB:AB002093; NID:92642444; PIDN:AAB87112.1; GSPDB:GN00139  
C/Genetics:  
A/Gene: T20D16.18; At2g23190  
A/Map position: 2  
A/Introns: 211/3; 339/3  
A/Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
A/Keywords: heme; iron; metalloprotein  
A/341-500/Domain: cytochrome P450 homology <P45>  
A/478/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 57.6%; Score 38; DB 2; Length 543;  
Best Local Similarity 83.3%; Pred. No. 34;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

XY 1 CFQWQR 6  
|||  
DB 498 CFQWQR 503

RESULT 14  
T14803  
phytochrome C - sorghum  
A/Species: Sorghum bicolor (sorghum)  
A/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 05-May-2000  
A/Accession: T14803  
A/Childs, K.L.; Miller, F.R.; Cordonnier-Pratt, M.M.; Pratt, L.H.; Morgan, P.W.; Mullet,  
submitted to the EMBL Data Library, April 1996  
A/Reference number: Z18186  
A/Accession: T14803  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-1135 <CHI>  
A/Cross-references: EMBL:U56731; NID:g1800218; PID:g1800219  
C/Genetics:  
A/Gene: PHYC  
A/Superfamily: phytochrome; phytochrome homology  
A/Keywords: chromoprotein; dimer; photoreceptor; phytochromobilin; transcription regulat  
A/65-581/Domain: phytochrome homology <PHY>  
A/321/Binding site: phytochromobilin (Cys) (covalent) #status predicted

Query Match 57.6%; Score 38; DB 2; Length 1135;  
Best Local Similarity 36.4%; Pred. No. 69;  
Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

XY 1 CFQWQRXMRKV 11  
|||  
DB 775 CLEWKNXMQKI 785

RESULT 15  
AG0794  
polymyxin B resistance protein [imported] - Salmonella enterica subsp. enterica serovar  
A/Species: Salmonella enterica subsp. enterica serovar Typhi  
A/Note: this species has also been called Salmonella typhi  
A/Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 09-Nov-2001  
A/Accession: AG0794  
A/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
h, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.  
A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
A/Reference number: AB0502; PMID:11677608  
A/Accession: AG0794  
A/Status: preliminary  
A/Molecule type: DNA

A/Residues: 1-85 <PAR>  
A/Cross-references: GB:AL513382; PIDN:CAD07537.1; PID:g16503529; GSPDB:GN00176  
C/Genetics:  
A/Gene: pmrD

Query Match 56.1%; Score 37; DB 2; Length 85;  
Best Local Similarity 60.0%; Pred. No. 8.9;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQRXMRK 10  
|||  
DB 74 CDEWQLTRK 83

Search completed: February 21, 2003, 08:02:45  
Job time: 11.6047 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

DM protein - protein search, using sw model

Run on: February 21, 2003, 07:28:06 ; Search time 5.2093 Seconds  
(without alignments)  
95.544 Million cell updates/sec

Title: US-09-743-107B-90  
Perfect score: 66  
Sequence: 1 CFQQRXMRKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	97.0	711	1 TRFL_HUMAN	P02788 homo sapien
2	53	80.3	708	1 TRFL_CAMDR	Q9tumo camelus dro
3	53	80.3	708	1 TRFL_CAPHI	Q29477 capra hircu
4	44	66.7	707	1 TRFL_MOUSE	P08071 mus musculu
5	43	65.2	4568	1 DYHB_CHLRE	Q39565 chlamydomon
6	39	59.1	695	1 TRFL_HORSE	Q77811 equus cabal
7	38	57.6	292	1 NLA_DROME	Q9x218 drosophila
8	38	57.6	1135	1 PHYC_SORBI	P93528 sorghum bic
9	37	56.1	85	1 PMRD_SALTY	P37589 salmonella
10	37	56.1	146	1 RPOB_HUMAN	P41187 liberibacte
11	36	54.5	235	1 FL3L_HUMAN	P49771 homo sapien
12	36	54.5	275	1 IL2A_BOVIN	P12342 bos taurus
13	36	54.5	275	1 IL2A_SHEEP	P26898 ovis aries
14	36	54.5	1137	1 PHYC_ORISA	Q92w19 oryza sativ
15	36	54.5	1179	1 ATX1_ARATH	Q91t02 arabidopsis
16	35	53.0	270	1 PDHX_MYXXA	P21159 myxococcus
17	35	53.0	306	1 BUB2_YEAST	P26448 saccharomyc
18	35	53.0	435	1 DCOR_PANRE	P49725 panagrellus
19	35	53.0	502	1 C911_ARATH	Q9f655 arabidopsis
20	35	53.0	783	1 YNR2_CAEEL	Q21388 caenorhabdi
21	35	53.0	2594	1 TLES_DROVI	P20806 drosophila
22	34	51.5	211	1 LOLB_VIBCH	P51070 vibrio chol
23	34	51.5	227	1 CHEZ_PSEAE	Q51434 pseudomonas
24	34	51.5	275	1 VA16_VACCV	P16710 vaccinia vi
25	34	51.5	378	1 VA16_VACCV	P20993 vaccinia vi
26	34	51.5	455	1 RADA_CHLMU	Q9pk96 chlamydia m
27	34	51.5	500	1 TLCE_RICPR	C00562 rickettsia
28	34	51.5	708	1 TRFL_EUBBU	Q77698 bubalus bub
29	34	51.5	765	1 Y008 HUNPO	Q15398 homo sapien
30	34	51.5	962	1 YBX7_SCHPO	Q10201 schizosacch
31	33	50.0	159	1 Y1CN_ECOLI	P31439 escherichia
32	33	50.0	272	1 IL2A_HUMAN	P01589 homo sapien
33	33	50.0	377	1 VA16_VARV	P33841 variola vir

34 33 50.0 428 1 SYH\_CHLMU Q9pij9 chlamydia m  
35 33 50.0 475 1 YEB1\_HAEIN P4493 haemophilus  
36 33 50.0 456 1 MSS1\_SCHPO Q9uce7 schizosacch  
37 33 50.0 506 1 AER\_ECOLI P50466 escherichia  
38 33 50.0 704 1 TRFL\_PIG P14632 sus scrofa  
39 33 50.0 1131 1 PHY\_PINSY Q41046 pinus sylve  
40 33 50.0 1213 1 T2D2\_DROME Q24325 drosophila  
41 33 50.0 1293 1 XPC\_DROME Q24595 drosophila  
42 33 50.0 2554 1 7LES\_DROME P13368 drosophila  
43 32 48.5 77 1 NULM\_ASCSU P24874 ascaris suu  
44 32 48.5 137 1 YEDD\_ECOLI P31063 escherichia  
45 32 48.5 141 1 YEDD\_SALTY Q06399 salmonella

## ALIGNMENTS

RESULT 1  
TRFL\_HUMAN  
ID TRFL\_HUMAN STANDARD; PRT; 711 AA.  
AC P02788; Q16780; Q16785; Q16786; Q16789; Q00756; Q9H1Z3; Q96KZ4;  
AC Q96KZ5;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Lactotransferrin precursor (Lactoferrin) [Contains: Lactoferrin A;  
DE Lactoferrin B; Lactoferrin C].  
GN LTF OR LF.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RX MEDLINE=90384839; PubMed=2402455;  
RA Rev M.W., Woloshuk S.L., de Boer H.A., Pieper F.R.;  
RT "Complete nucleotide sequence of human mammary gland lactoferrin.";  
RL Nucleic Acids Res. 18:5288-5288(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RX Cho Y.Y.;  
RL Thesis (1994), Genetic Engineering Research Institute / Taejon, Korea.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Connely O.M.;  
RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RX Liang Q., Jimenez-Flores R., Richardson T.;  
RT "Molecular cloning and sequence analysis of human lactoferrin.";  
RL Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Bone marrow;  
RA Wei X., Han J., Rado T.A.;  
RT "Human neutrophil lactoferrin coding and 5' flanking region DNA  
sequences.";  
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RX Cheng H., Chen X., Huan L.;  
RT "cDNA cloning and sequence analysis of human lactoferrin.";  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Prostate;  
RA Strausberg R.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
RN [8]

RP SEQUENCE OF 3-711 FROM N.A.  
RC TISSUE-Mammary gland.  
RX MEDLINE=90326549; PubMed=2374734;  
RA Powell M.J., Oden J.B.;  
RT "Nucleotide sequence of human lactoferrin cDNA.";  
RL Nucleic Acids Res. 18:4013-4013(1990).  
RN [9]  
RP SEQUENCE OF 20-711.  
RX MEDLINE=85076667; PubMed=6510420;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Schoentgen F.,  
RT Metz-Boutigue M.-H., Jolles J., Jolles P.;  
RL "Human lactoferrin: amino acid sequence and structural  
RT comparisons with other transferrins.";  
RL Eur. J. Biochem. 145:659-666(1984).  
RN [10]  
RP PRELIMINARY SEQUENCE OF 20-73; 134-171; 257-278 AND 347-530.  
RX MEDLINE=82046817; PubMed=6794640;  
RA Metz-Boutigue M.-H., Mazurier J., Jolles J., Spik G., Montreuil J.,  
RT Jolles P.;  
RL "The present state of the human lactotransferrin sequence. Study and  
RT alignment of the cyanogen bromide fragments and characterization of  
RT N- and C-terminal domains.";  
RL Biochim. Biophys. Acta 670:243-254(1981).  
RN [11]  
RP SEQUENCE OF 609-711.  
RX MEDLINE=82262043; PubMed=7049727;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,  
RT Jolles P.;  
RL "An 88 amino acid long C-terminal sequence of human  
RT lactotransferrin.";  
RL FEBS Lett. 142:107-110(1982).  
RN [12]  
RP SEQUENCE OF 436-711 FROM N.A.  
RX MEDLINE=88001031; PubMed=3477300;  
RA Rado T.A., Wei X., Benz E.J. Jr.;  
RT "Isolation of lactoferrin cDNA from a human myeloid library and  
RT expression of mRNA during normal and leukemic myelopoiesis.";  
RL Blood 70:989-993(1987).  
RN [13]  
RP SEQUENCE OF 237-711 FROM N.A.  
RA McCombie W.R., Wilson R., Chen B., Gibbs R., Zuo L., Johnson D.,  
RT Nho M., Parnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,  
RA Gao L., la Bastide M., Kaplan N., Greco T., Touchman J., Muzny D.,  
RT Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M., Porcel B.M.,  
RA Dragan Y., Giacalone J., Pae A., Powell E., Solinsky K.A., Desilva U.,  
RA Diaz-Perez S., Zhou X., Yu Y., Watanabe M., Doggett N., Garcia D.,  
RA Sadrilanti J.L.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
RN [14]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND REVISIONS.  
RX MEDLINE=90064528; PubMed=2585506;  
RA Anderson B.F., Baker H.M., Norris G.E., Rice D.W., Baker E.N.;  
RT "Structure of human lactoferrin: crystallographic structure analysis  
RT and refinement at 2.8-A resolution.";  
RL J. Mol. Biol. 209:711-734(1989).  
RN [15]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RA Haridas M., Anderson B.F., Baker E.N.;  
RT "Structure of human diferric lactoferrin refined at 2.2-A  
RT resolution.";  
RL Acta Crystallogr. D 51:629-646(1995).  
RN [16]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 20-353.  
RX MEDLINE=97156796; PubMed=9003186;  
RA Nicholson H., Anderson B.F., Bland T., Shewry S.C., Tweedie J.W.,  
RT Baker E.N.;  
RL "Mutagenesis of the histidine ligand in human lactoferrin: iron  
RT binding properties and crystal structure of the histidine-  
RT 253->-methionine mutant.";  
RL Biochemistry 36:341-346(1997).  
RN [17]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX MEDLINE=99190892; PubMed=10089347;  
RA Sun X.L., Baker H.M., Shewry S.C., Jameson G.B., Baker E.N.;  
RT "Structure of recombinant human lactoferrin expressed in Aspergillus  
RT avamori.";  
RL Acta Crystallogr. D 55:403-407(1999).  
RN [18]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=99192677; PubMed=10089508;  
RA Jameson G.B., Anderson B.F., Norris G.E., Thomas D.H., Baker E.N.;  
RT "Structure of human apolactoferrin at 2.0-A resolution. Refinement  
RT and analysis of ligand-induced conformational change.";  
RL Acta Crystallogr. D 54:1319-1335(1998).  
RN [19]  
RP CHARACTERIZATION OF LACTOFERROXINS.  
RX MEDLINE=91166929; PubMed=1369293;  
RA Tani F., Iio K., Chiba H., Yoshikawa M.;  
RT "Isolation and characterization of opioic antagonist peptides derived  
RT from human lactoferrin.";  
RL Agric. Biol. Chem. 54:1803-1810(1990).  
RN [20]  
RP VARIANTS THR-30 AND ARG-48.  
RX PubMed=9873069;  
RA Klintworth G.K., Sommer J.R., Obrian G., Han L., Ahmed M.N.,  
RT Qumsiyeh M.B., Lin P.-Y., Basti S., Reddy M.K., Kanai A., Hotta Y.,  
RA Sugar J., Kumaranickavel G., Munier F., Schorderet D.P.,  
RT El Matiri L., Iwata F., Kaiser-Kupfer M., Negata M., Nakayasu K.,  
RA Hejtmancik J.F., Teng C.T.;  
RT "Familial subepithelial corneal amyloidosis (gelatinous drop-like  
RT corneal dystrophy): exclusion of linkage to lactoferrin gene.";  
RL Mol. Vision 4:31-32(1998).  
RN [21]  
RP FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
CC OF AN ANION, USUALLY BICARBONATE.  
CC [22]  
RP FUNCTION: LACTOFERROXIN A, B AND C HAVE OPIOID ANTAGONIST  
CC ACTIVITY. LACTOFERROXIN A SHOWS PREFERENCE FOR MU-RECEPTORS, WHILE  
CC LACTOFERROXIN B AND C HAVE SOMEWHAT HIGHER DEGREES OF PREFERENCE  
CC FOR KAPPA-RECEPTORS THAN FOR MU-RECEPTORS.  
CC [23]  
RP SUBUNIT: MONOMER.  
CC [24]  
RP SUBCELLULAR LOCATION: Secreted.  
CC [25]  
RP DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
CC [26]  
RP SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
CC [27]  
RP THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC [28]  
RP EMBL; X53961; CAA37914.1; -  
DR EMBL; U07643; AAB60324.1; -  
DR EMBL; M91150; AAA36159.1; -  
DR EMBL; M83202; AAA59511.1; -  
DR EMBL; M83205; AAA86656.1; -  
DR EMBL; M18642; AAA86655.1; -  
DR EMBL; AF332168; AAG48753.1; -  
DR EMBL; BC015822; AAH15822.1; -  
DR EMBL; BC015823; AAH15823.1; -  
DR EMBL; M73700; AAA59479.1; -  
DR EMBL; X52941; CAA37116.1; -  
DR EMBL; U95626; AAB57795.1; -  
DR PIR; S11228; TPFHUL.  
DR PDB; 1LCF; 31-AUG-94.  
DR PDB; 1LCT; 31-OCT-93.  
DR PDB; 1LFG; 31-JUL-94.  
DR PDB; 1LFH; 31-OCT-93.  
DR PDB; 1LFI; 31-OCT-93.  
DR PDB; 1LGB; 31-AUG-94.  
DR PDB; 1LGC; 31-AUG-94.  
DR PDB; 1LKA; 08-NOV-96.  
DR PDB; 1DSN; 08-MAR-96.  
DR PDB; 1HSE; 12-MAR-97.  
DR PDB; 1VFD; 21-APR-97.

```

Query Match          97.0%; Score 64; DB 1; Length 711;
Best Local Similarity 91.7%; Pred. NO. 0.00024;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQRXMKVR 12
DB 39 CFQWQRXMKVR 50

RESULT 2
ID TRFL_CAMDR STANDARD; PRT; 708 AA.
AC Q9TUM0; Q9MZS5;
JT 16-OCT-2001 (Rel. 40, Created)
JT 16-OCT-2001 (Rel. 40, Last sequence update)
JT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lactotransferrin precursor (Lactoferrin).
EN Lf.

DS Camelus dromedarius (Dromedary) (Arabic camel).
DC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.
DX NCBI_TaxID=9838;
ZN [1]
ZP SEQUENCE FROM N.A.
ZC STRAIN=Somali; TISSUE=Lactating mammary gland;
ZA Kappeler S.R.; Ackermann M.; Farah Z.; Puhan Z.;
ZT "Sequence analysis of camel (Camelus dromedarius) lactoferrin.";
ZL Int. Dairy J. 9:481-486(1999).
ZN [2]
ZP SEQUENCE FROM N.A.
ZC TISSUE=Mammary gland;
ZA Paramasivam M., Srinivasan A., Singh R., Sahani M.S., Singh T.P.;
ZL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
DC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
DC OF AN ANION, USUALLY BICARBONATE (BY SIMILARITY).
DC -!- SUBUNIT: MONOMER (BY SIMILARITY).
DC -!- SUBCELLULAR LOCATION: Secreted.
DC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
DC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
DC -----
DC This SWISS-PROT entry is copyright. It is produced through a collaboration
DC between the Swiss Institute of Bioinformatics and the EMBL outstation -
DC the European Bioinformatics Institute. There are no restrictions on its
DC use by non-profit institutions as long as its content is in no way
DC modified and this statement is not removed. Usage by and for commercial
DC entities requires a license agreement (See http://www.isb-sib.ch/announce/
DC or send an email to license@isb-sib.ch).
DC -----
DR EMBL; AJ131674; CAB53387.1; -
DR EMBL; AF165879; AAF82241.1; -
DR HSRP; O77811; IBIX.
DR InterPro; IPR011156; Transferrin.
DR Pfam; PF00405; transferrin; 2.
DR PRINTS; PR00422; TRANSFERRIN.
DR SMART; SM00094; TR_FER; 2.
DR PROSITE; PS00205; TRANSFERRIN 1; 2.
DR PROSITE; PS00206; TRANSFERRIN 2; 2.
DR PROSITE; PS00207; TRANSFERRIN 3; 2.
CW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;
CW Signal.
PT CHAIN 1 19 BY SIMILARITY.
PT REPEAT 20 708 LACTOTRANSFERRIN.
PT REPEAT 20 363 1.
PT REPEAT 364 708 2.
PT DISULFID 28 64 BY SIMILARITY.
PT DISULFID 38 55 BY SIMILARITY.
PT DISULFID 134 217 BY SIMILARITY.
PT DISULFID 176 192 BY SIMILARITY.
PT DISULFID 189 200 BY SIMILARITY.
PT DISULFID 250 264 BY SIMILARITY.
PT DISULFID 367 399 BY SIMILARITY.

```

```

FT DISULFID 377 390 BY SIMILARITY.
FT DISULFID 424 703 BY SIMILARITY.
FT DISULFID 444 666 BY SIMILARITY.
FT DISULFID 476 551 BY SIMILARITY.
FT DISULFID 500 694 BY SIMILARITY.
FT DISULFID 510 524 BY SIMILARITY.
FT DISULFID 521 534 BY SIMILARITY.
FT DISULFID 592 606 BY SIMILARITY.
FT DISULFID 644 649 BY SIMILARITY.
FT METAL 79 IRON 1 (BY SIMILARITY).
FT METAL 111 IRON 1 (BY SIMILARITY).
FT METAL 211 IRON 1 (BY SIMILARITY).
FT METAL 272 272 IRON 1 (BY SIMILARITY).
FT METAL 414 414 IRON 2 (BY SIMILARITY).
FT METAL 452 452 IRON 2 (BY SIMILARITY).
FT METAL 545 545 IRON 2 (BY SIMILARITY).
FT METAL 614 614 IRON 2 (BY SIMILARITY).
FT BINDING 140 140 ANION (BY SIMILARITY).
FT BINDING 482 482 ANION (BY SIMILARITY).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 385 385 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 537 537 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 594 594 F -> S (IN REF. 2).
FT CONFLICT 261 261 G -> A (IN REF. 2).
FT CONFLICT 304 304 S -> P (IN REF. 2).
FT CONFLICT 330 330 LLS -> PLF (IN REF. 2).
FT CONFLICT 492 494 L -> F (IN REF. 2).
FT CONFLICT 506 506 A -> P (IN REF. 2).
FT CONFLICT 609 609 R -> Q (IN REF. 2).
FT CONFLICT 642 642
SQ SEQUENCE 708 AA; 77211 MW; 0B0C175A0B69D430 CRC64;

Query Match          80.3%; Score 53; DB 1; Length 708;
Best Local Similarity 75.0%; Pred. NO. 0.025;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQRXMKVR 12
DB 38 CAQWQRXMKVR 49

RESULT 3
ID TRFL_CAPHI STANDARD; PRT; 708 AA.
AC Q29477; Q29479;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lactotransferrin precursor (Lactoferrin).
EN Lf.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Lee T., Yu S., Kim S., Lee K., Yu D.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=94380047; PubMed=8093048;
RA Le Provost F., Nocart M., Guerin G., Martin P.;
RT "Characterization of the goat lactoferrin cDNA. Assignment of the
RL relevant locus to bovine U12 syntenic group.";
RL Biochem. Biophys. Res. Commun. 203:1324-1332(1994).
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.

```







-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
-----

EMBL; AJ010930; CAA09407.1; --  
PDB; 1BIX; 02-DEC-98.  
PDB; 1B7U; 02-FEB-99.  
PDB; 1B7Z; 02-FEB-99.  
InterPro; IPR001156; Transferrin.  
Pfam; PF00405; transferrin; 2.  
PRINTS; PR00422; TRANSFERRIN.  
SMART; SM00094; TR\_FER; 2.  
PROSITE; PS00205; TRANSFERRIN\_1; 2.  
PROSITE; PS00206; TRANSFERRIN\_2; 2.  
PROSITE; PS00207; TRANSFERRIN\_3; 1.  
KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
Signal; 3D-structure.  
FT SIGNAL 1  
FT CHAIN <1 6 LACTOTRANSFERRIN.  
FT REPEAT 7 695  
FT REPEAT 351 695  
FT DISULFID 15 51  
FT DISULFID 25 42  
FT DISULFID 121 204  
FT DISULFID 163 179  
FT DISULFID 186 189  
FT DISULFID 176 187  
FT DISULFID 237 251  
FT DISULFID 354 386  
FT DISULFID 364 377  
FT DISULFID 411 690  
FT DISULFID 431 653  
FT DISULFID 463 538  
FT DISULFID 487 691  
FT DISULFID 497 511  
FT DISULFID 508 521  
FT DISULFID 579 593  
FT DISULFID 631 636  
FT METAL 66 56  
FT METAL 98 98  
FT METAL 198 198  
FT METAL 259 259  
FT METAL 401 401  
FT METAL 439 439  
FT METAL 532 532  
FT METAL 601 601  
FT ANION (BY SIMILARITY).  
FT ANION (BY SIMILARITY).  
FT ANION (BY SIMILARITY).  
FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 287 287  
FT CARBOHYD 482 482  
SQ SEQUENCE 695 AA; 75991 MW; 07BB84D50E1B165D CRC64;

Query Match 59.1%; Score 39; DB 1; Length 695;  
Best Local Similarity 58.3%; Pred. No. 9.3;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CFQWQRXMKVR 12  
| : : : : :  
Db 25 CAKFRNMKVR 36  
RESULT 7  
ID NLA DROME  
AC Q9XZL8; Q9V391; PRT; 292 AA.  
DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Nebula protein.  
GN NLA OR CG6072.  
OS Drosophila melanogaster (Fruit fly)  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA McCormick A.V., Goldberg M.L.;  
RT "Gene required for elongation of meiosis I spindle in Drosophila females."  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkley;  
RX MEDLINE=2019606; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Calle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busam D.A., Bultke H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,  
RA Fostler C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpene G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko F., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mout S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Wang X.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang Z.-Y.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster."  
RL Science 287:2185-2195(2000).  
CC -!- FUNCTION: REQUIRED FOR ELONGATION OF MEIOSIS I SPINDLE.  
CC -!- SIMILARITY: BELONGS TO THE DSCRI FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
-----  
EMBL; AF147700; AAD33987.1; -  
EMBL; AE003712; AAF55285.1; -  
DR FlyBase; Fegno0026629; nla.  
SQ SEQUENCE 292 AA; 31423 MW; 64F1BBF5F6EA6CF9 CRC64;

Query Match 57.6%; Score 38; DB 1; Length 292;  
 Best Local Similarity 54.5%; Pred. No. 6;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Y 2 FQWQXMRKV 12  
 |||||:  
 b 150 FQWLSFRLR 160

RESULT 8  
 HYC\_SORBI  
 D -HYC\_SORBI STANDARD; PRT; 1135 AA.  
 C P93528;  
 Y 16-OCT-2001 (Rel. 40, Created)  
 Y 16-OCT-2001 (Rel. 40, Last sequence update)  
 Y 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Phytochrome C.  
 N PHYC.  
 S Sorghum bicolor (Sorghum) (Sorghum vulgare).  
 C Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 C Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
 C Panicoideae; Andropogoneae; Sorghum.  
 X NCBI\_TaxID=4558;  
 [1]  
 SEQUENCE FROM N.A.  
 P MEDLINE=97198556; PubMed=9046599;  
 A Childs K.L., Miller F.R., Cordonnier-Pratt M.-M., Pratt L.H.,  
 A Morgan P.W., Mullet J.E.;  
 T "The Sorghum bicolor photoperiod sensitivity gene, Ma3, encodes a  
 T phytochrome B.";  
 L Plant Physiol. 113:611-619 (1997).  
 C -!- FUNCTION: REGULATORY PHOTORECEPTOR WHICH EXISTS IN TWO FORMS THAT  
 C ARE REVERSIBLY INTERCONVERTIBLE BY LIGHT: THE PR FORM THAT ABSORBS  
 C MAXIMALLY IN THE RED REGION OF THE SPECTRUM AND THE PFR FORM THAT  
 C ABSORBS MAXIMALLY IN THE FAR-RED REGION. PHOTOCONVERSION OF PR IN  
 C PFR INDUCES AN ARRAY OF MORPHOGENIC RESPONSES, WHEREAS  
 C RECONVERSION OF PFR TO PR CANCELS THE INDUCTION OF THOSE  
 C RESPONSES. PFR CONTROLS THE EXPRESSION OF A NUMBER OF NUCLEAR  
 C GENES INCLUDING THOSE ENCODING THE SMALL SUBUNIT OF RIBULOSE-  
 C BISPHOSPHATE CARBOXYLASE, CHLOROPHYLL A/B BINDING PROTEIN,  
 C PROTOCHLOROPHYLLIDE REDUCTASE, RNA, ETC. IT ALSO CONTROLS THE  
 C EXPRESSION OF ITS OWN GENE(S) IN A NEGATIVE FEEDBACK FASHION (BY  
 C SIMILARITY).  
 C -!- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 C -!- PTM: CONTAINS ONE COVALENTLY LINKED TETRAPYRROLE CHROMOPHORE.  
 C -!- SIMILARITY: BELONGS TO THE PHYTOCHROME FAMILY.  
 C -!- SIMILARITY: CONTAINS 2 PAS (PER-ARNT-SIM) DIMERIZATION DOMAINS.  
 C -!- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.  
 C -!- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.  
 C -----  
 C This SWISS-PROT entry is copyright. It is produced through a collaboration  
 C between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 C the European Bioinformatics Institute. There are no restrictions on its  
 C use by non-profit institutions as long as its content is in no way  
 C modified and this statement is not removed. Usage by and for commercial  
 C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 C or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 C -----  
 C EMBL; U56731; AAB41399.1;  
 C InterPro; IPR003594; ATPbind\_ATPase.  
 C InterPro; IPR003018; GAF.  
 C InterPro; IPR004359; HIS\_KIN\_sig.  
 C InterPro; IPR003661; HIS\_KIN.  
 C InterPro; IPR001610; PAC.  
 C InterPro; IPR000014; PAS\_domain.  
 C InterPro; IPR001294; Phytochrome.  
 C Pfam; PF00360; phytochrome; 1.  
 C Pfam; PF00512; signal; 1.  
 C Pfam; PF00989; PAS; 2.  
 C Pfam; PF01590; GAF; 1.  
 C Pfam; PF02519; HATPase\_c; 1.  
 C PRINTS; PR01033; PHYTOCHROME.  
 C SMART; SM00065; GAF; 1.

DR SMART; SM00387; HATPase\_c; 1.  
 DR SMART; SM00388; HSKA; 1.  
 DR SMART; SM00086; PAC; 1.  
 DR SMART; SM00091; PAS; 2.  
 DR TIGRFS; TIGR00229; sensory\_box; 2.  
 DR PROSITE; PSS0109; HIS\_KIN; 1.  
 DR PROSITE; PSS0112; PAS; 2.  
 DR PROSITE; PSS0245; PHYTOCHROME\_1; FALSE\_NEG.  
 DR PROSITE; PSS0046; PHYTOCHROME\_2; 1.  
 KW Transcription regulation; Photoreceptor; Phytochrome; Chromophore;  
 KW Repeat; Multigene family.  
 FT DOMAIN 618 688 PAS 1.  
 FT DOMAIN 748 822 PAS 2.  
 FT DOMAIN 902 1122 HISTIDINE KINASE.  
 FT BINDING 321 321 CHROMOPHORE (BY SIMILARITY).  
 SQ SEQUENCE 1135 AA; 126315 MW; AFCC934B7592DE4D CRC64;

Query Match 57.6%; Score 38; DB 1; Length 1135;  
 Best Local Similarity 36.4%; Pred. No. 23;  
 Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Oy 1 CFQWQXMRKV 11  
 |:::  
 Db 775 CLEWKAQKI 785

RESULT 9  
 PMRD\_SALTY  
 ID PMRD\_SALTY STANDARD; PRT; 85 AA.  
 AC P37589;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Polymyxin B resistance protein pmrD.  
 GN PMRD OR STM2304.  
 OS Salmonella typhimurium.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Salmonella.  
 OX NCBI\_TaxID=602;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=J2;  
 RX MEDLINE=94266712; PubMed=8206837;  
 RA Roland K.L., Esther C.R., Spitznagel J.K.;  
 RT "Isolation and characterization of a gene, pmrD, from Salmonella  
 RT typhimurium that confers resistance to polymyxin when expressed in  
 RT multiple copies.";  
 RL J. Bacteriol. 176:3589-3597 (1994).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=J2 / SGSC1412 / ATCC 700720;  
 RX MEDLINE=21534948; PubMed=11677609;  
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
 RA Waterston R., Wilson R.K.;  
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
 RT J2.";  
 RL Nature 413:852-856 (2001).  
 CC -!- FUNCTION: CONFERS RESISTANCE TO POLYMYXIN B. POLYMYXIN RESISTANCE  
 CC MAY BE MEDIATED BY AN INTERACTION BETWEEN PMRA OR A PMRA-REGULATED  
 CC GENE PRODUCT AND PMRD.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U02281; AAA21322.1; -.

DR EMBL; AE008803; AAL21205.1; --  
 DR StryGene; SGI0304; pmd. Complete proteome.  
 KW Antibiotic resistance; 85 AA; 9749 MW; 1B1922419EA50CCA CRC64;  
 SQ SEQUENCE 85 AA; 9749 MW; 1B1922419EA50CCA CRC64;  
 Query Match 56.1%; Score 37; DB 1; Length 85;  
 Best Local Similarity 60.0%; Pred. No. 2.7;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 CFOWQXMRK 10  
 Db 74 CDEWQLTRX 83  
 RESULT 10  
 RPOB LIBAF STANDARD; PRT; 146 AA.  
 ID RPOB LIBAF STANDARD; PRT; 146 AA.  
 AC P41187;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE DNA-directed RNA polymerase beta chain [EC 2.7.7.6] (Transcriptase  
 DE beta chain) (RNA polymerase beta subunit) (Fragment).  
 GN RPOB.  
 OS Liberibacter africanus (Liberibacter africanus).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Candidatus Liberibacter.  
 OX NCBI\_TaxID=34020;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Nelspruit;  
 RA Planet P., Jagueix S., Bove J.M., Garnier M.;  
 RT "Detection and characterization of the African Citrus Greening  
 RT Liberibacter by amplification, cloning and sequencing of the rplKAL-  
 RT rpoBC operon";  
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION  
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS  
 CC SUBSTRATES.  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC {RNA}(N).  
 CC -!- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE  
 CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1  
 CC BETA' CHAIN.  
 CC -!- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U09675; AAA19557.1; --  
 CC InterPro; IPR001572; RNA pol B.  
 CC Pfam; PF00562; RNA pol B.1.  
 CC PROSITE; PS01166; RNA POL BETA, PARTIAL.  
 KW Transferase; transcription; DNA-directed RNA polymerase.  
 FT NON TER 146  
 SQ SEQUENCE 146 AA; 16731 MW; 815D6EDD6FD09943 CRC64;  
 Query Match 56.1%; Score 37; DB 1; Length 146;  
 Best Local Similarity 60.0%; Pred. No. 4.6;  
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 CFOWQXMRK 10  
 Db 10 CVOWSGARK 19  
 RESULT 11  
 FL3L\_HUMAN

ID AC  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE SL cytokine precursor (Fms-related tyrosine kinase 3 ligand) (Flt3  
 DE ligand) (Flt3l).  
 GN FLT3LG.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94195428; PubMed=8145851;  
 RA Hannum C., Culpepper J., Campbell D., McClanahan T., Zurawski S.,  
 RA Bada J.F., Kastelein R., Hudak S., Wagner J., Mattson J., Luh J.,  
 RA Duda G., Martina N., Peterson D., Menon S., Shanafelt A.,  
 RA Muench M., Kelnner G., Namikawa R., Rennick D., Roncarolo M.G.,  
 RA Zlotnik A., Rosner O., Dubreuil P., Birnbaum D., Lee F.;  
 RT "Ligand for FLT3/FLK2 receptor tyrosine kinase regulates growth of  
 RT haematopoietic stem cells and is encoded by variant RNAs";  
 RL Nature 368:643-648(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94235842; PubMed=8180375;  
 RA Lyman S.D., James L., Johnson L., Brasel K., de Vries P.,  
 RA Escobar S.S., Downey H., Spelt R.R., Beckmann M.P., McKenna H.J.;  
 RT "Cloning of the human homologue of the murine flt3 ligand: a growth  
 RT factor for early hematopoietic progenitor cells";  
 RL Blood 83:2795-2801(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RX MEDLINE=96032581; PubMed=7566977;  
 RA Lyman S.D., Stocking K., Davison B., Fletcher F., Johnson L.,  
 RA Escobar S.;  
 RT "Structural analysis of human and murine flt3 ligand genomic loci";  
 RL Oncogene 11:1165-1172(1995).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RX MEDLINE=20343011; PubMed=1081197;  
 RA Savvides S.N., Boone T., Karplus P.A.;  
 RT "Flt3 ligand structure and unexpected commonalities of helical  
 RT bundles and cysteine knots";  
 RL Nat. Struct. Biol. 7:486-491(2000).  
 CC -!- FUNCTION: STIMULATES THE PROLIFERATION OF EARLY HEMATOPOIETIC  
 CC CELLS. SYNERGIZES WELL WITH A NUMBER OF OTHER COLONY STIMULATING  
 CC FACTORS AND INTERLEUKINS.  
 CC -!- SUBUNIT: Homodimer (isoform 2).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1);  
 CC secreted (isoform 2).  
 CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; a 1/membrane-bound (shown here)  
 CC and 2/soluble; are produced by alternative splicing.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U04806; AAA17999.1; --  
 CC EMBL; U03858; AAA19825.1; --  
 CC EMBL; U29874; AAA90949.1; --  
 CC EMBL; U29874; AAA90950.1; --  
 CC PDB; 1ETE; 09-JUN-00.  
 CC Genew; HGNC:3766; FLT3LG.  
 CC MIM; 600007; --  
 CC InterPro; IPR004213; Flt3\_lig.  
 CC Pfam; PF02947; flt3\_lig.1.  
 KW Cytokine; Glycoprotein; Transmembrane; Alternative splicing; Signal;  
 KW 3D-structure.

```
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 235 SL CYTOKINE.
FT DOMAIN 27 184 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 185 205 POTENTIAL.
FT DOMAIN 206 235 CYTOPLASMIC (POTENTIAL).
FT DISULFID 30 111
FT DISULFID 30 153
FT DISULFID 119 158
FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 161 178 DSSTLPWPWPGRPLEATA -> VETVPHRVSDQGLDLS
FT VARSPLIC 179 235 (IN ISOFORM 2).
FT VARSPLIC 179 235 MISSING (IN ISOFORM 2).
FT CONFLICT 72 72 G -> A (IN REF. 1).
FT CONFLICT 72 72
SQ SEQUENCE 235 AA; 26416 MW; 73895BF693B4CECF CRC64;

Query Match 54.5%; Score 36; DB 1; Length 235;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

2Y 1 CFQWQRYXMK 10
3B 204 CLHWQRTERR 213

RESULT 12
IL2A_BOVIN STANDARD; PRT; 275 AA.
AC P12342;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Interleukin-2 receptor alpha chain precursor (IL-2 receptor alpha
DE subunit) (P55) (TAC antigen) (CD25).
IL2RA.
3N Bos taurus (Bovine).
3S Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
3C Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
3C Bovidae; Bovinae; Bos.
3X NCBI_TaxID=9913;
3N [1]_TaxID=9913;
3N [2]_TaxID=9913;
3N SEQUENCE FROM N.A.
3R MEDLINE=8212503; PubMed=2835311;
3A Weinberg A.D., Shaw J., Paetkau V., Bleackley R.C., Magnuson N.S.,
3A Reeves R., Magnuson J.A.;
3R "Cloning of cDNA for the bovine IL-2 receptor (bovine Tac antigen).";
3L Immunology 63:603-610(1988).
3N [2]
3R SEQUENCE OF 1-21 FROM N.A.
3R MEDLINE=96116968; PubMed=8563178;
3A Yoo J., de Leon F.A., Stone R.T., Beattie C.W.;
3R "Cloning and chromosomal assignment of the bovine interleukin-2
3R receptor alpha (IL-2R alpha) gene.";
3L Mamm. Genome 6:751-753(1995).
3C -!- FUNCTION: RECEPTOR FOR INTERLEUKIN-2.
3C -!- SUBUNIT: NON COVALENT DIMER OF AN ALPHA AND A BETA CHAINS. IL-2R
3C EXIST IN 3 DIFFERENT FORMS: A HIGH AFFINITY DIMER, AN INTERMEDIATE
3C AFFINITY MONOMER (BETA CHAIN), AND A LOW AFFINITY MONOMER (ALPHA
3C CHAIN). THE HIGH AND INTERMEDIATE AFFINITY FORMS ALSO ASSOCIATE
3C WITH A GAMMA CHAIN.
3C -!- SUBCELLULAR LOCATION: Type I membrane protein.
3C -!- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
3C -----
3C This SWISS-PROT entry is copyright. It is produced through a collaboration
3C between the Swiss Institute of Bioinformatics and the EMBL outstation
3C the European Bioinformatics Institute. There are no restrictions on its
3C use by non-profit institutions as long as its content is in no way
3C modified and this statement is not removed. Usage by and for commercial
3C entities requires a license agreement (See http://www.isb-sib.ch/announce/
3C or send an email to license@isb-sib.ch).
3C -----
3R EMBL; M20818; AAA51414.1; --
3R EMBL; U24226; AAC48487.1; --

DR PIR; S07442; S07442.
DR HSSP; P01569; IILM.
DR InterPro; IPR000436; Sush1_SCR_CCP.
DR Pfam; PF00084; sush1; 2.
DR SMART; SM00032; CCP; 2.
KW Transmembrane; Glycoprotein; Receptor; Repeat; Signal; Sushi.
FT SIGNAL 1 21
FT CHAIN 22 275 INTERLEUKIN-2 RECEPTOR ALPHA CHAIN.
FT DOMAIN 22 243 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 244 262 POTENTIAL.
FT DOMAIN 263 275 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 23 78 SUSHI 1.
FT DOMAIN 122 185 SUSHI 2.
FT DISULFID 24 64 BY SIMILARITY.
FT DISULFID 751 77 BY SIMILARITY.
FT DISULFID 123 168 BY SIMILARITY.
FT DISULFID 152 184 BY SIMILARITY.
FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 275 AA; 31238 MW; 4901BBF9A4862390 CRC64;

Query Match 54.5%; Score 36; DB 1; Length 275;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CFQWQRYXMKVR 12
DB 261 CLTWQKWKQVR 272

RESULT 13
IL2A_SHEEP STANDARD; PRT; 275 AA.
AC P26898;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Interleukin-2 receptor alpha chain precursor (IL-2 receptor alpha
DE subunit) (P55) (TAC antigen) (CD25).
IL2RA.
3S Ovis aries (Sheep).
3C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
3C Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
3C Bovidae; Caprinae; Ovis.
3X NCBI_TaxID=9940;
3N [1]_TaxID=9940;
3N [2]_TaxID=9940;
3R SEQUENCE FROM N.A.
3R TISSUE=T-cell;
3A Verhagen A.A.;
3R Submitted (DEC-1991) to the EMBL/GenBank/DBSJ databases.
3N [2]
3R SEQUENCE FROM N.A.
3R MEDLINE=92241682; PubMed=1572550;
3X Bujdosó R., Sargan D.R., Williamson M.L., McConnell I.;
3R "Cloning of a cDNA encoding the ovine interleukin-2 receptor 55-kDa
3R protein, CD25.";
3L Gene 113:283-284(1992).
3C -!- FUNCTION: RECEPTOR FOR INTERLEUKIN-2.
3C -!- SUBUNIT: NON COVALENT DIMER OF AN ALPHA AND A BETA CHAINS. IL-2R
3C EXIST IN 3 DIFFERENT FORMS: A HIGH AFFINITY DIMER, AN INTERMEDIATE
3C AFFINITY MONOMER (BETA CHAIN), AND A LOW AFFINITY MONOMER (ALPHA
3C CHAIN). THE HIGH AND INTERMEDIATE AFFINITY FORMS ALSO ASSOCIATE
3C WITH A GAMMA CHAIN.
3C -!- SUBCELLULAR LOCATION: Type I membrane protein.
3C -!- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
3C -----
3C This SWISS-PROT entry is copyright. It is produced through a collaboration
3C between the Swiss Institute of Bioinformatics and the EMBL outstation
3C the European Bioinformatics Institute. There are no restrictions on its
3C use by non-profit institutions as long as its content is in no way
3C modified and this statement is not removed. Usage by and for commercial
3C entities requires a license agreement (See http://www.isb-sib.ch/announce/
3C or send an email to license@isb-sib.ch).
3C -----
3R EMBL; M20818; AAA51414.1; --
3R EMBL; U24226; AAC48487.1; --
```

```

CC -----
DR ENBL; Z11560; CAA77652.1; -
DR ENBL; X60149; CAA42723.1; -
DR ENBL; A39167; CAA01447.1; -
DR PIR; S18910; S18910.
DR PIR; S18899; S18899.
DR PIR; JCI113; JCI113.
DR HSP; P01589; IILM.
DR InterPro; IPR000436; Sushi_scr_CCP.
DR Pfam; PF00084; sushi; 2.
DR SMART; SM00332; CCP; 2.
KW Transmembrane; Glycoprotein; Receptor; Repeat; Signal; Sushi.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 275 INTERLEUKIN-2 RECEPTOR ALPHA CHAIN.
FT DOMAIN 22 243 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 244 262 POTENTIAL.
FT DOMAIN 263 275 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 23 78 SUSHI 1.
FT DOMAIN 122 185 SUSHI 2.
FT DISULFID 24 64 BY SIMILARITY.
FT DISULFID 751 77 BY SIMILARITY.
FT DISULFID 123 168 BY SIMILARITY.
FT DISULFID 152 184 BY SIMILARITY.
FT CARBOHYD 180 180 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CONFLICT 186 186 S -> T (IN REF. 2).
SQ SEQUENCE 275 AA; 30904 MW; 1101A2DESACS5A088 CRC64;

Query Match 54.5%; Score 36; DB 1; Length 275;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12
DB 261 CLTWQRMKVR 272

RESULT 14
PHYC_ORYSA
ID PHYC_ORYSA STANDARD; PRT; 1137 AA.
AC Q9ZWI9; P93429;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phytochrome C.
GN PHYC.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Tahir M., Kanegae H., Takano M.;
RT "Phytochrome C (PHYC) gene in rice: isolation and characterization of
a complete coding sequence.";
RL (in) Plant Gene Register PGR98-210.
RN [2]
RP SEQUENCE OF 275-378 FROM N.A.
RX MEDLINE=97019052; PubMed=8865688;
RA Mathews S., Sharrock R.A.;
RT "The phytochrome gene family in grasses (Poaceae): a phylogeny and
evidence that grasses have a subset of the loci found in dicot
angiosperms.";
RL Mol. Biol. Evol. 13:1141-1150(1996).
CC -1- FUNCTION: REGULATORY PHOTORECEPTOR WHICH EXISTS IN TWO FORMS THAT
ARE REVERSIBLY INTERCONVERTIBLE BY LIGHT: THE PR FORM THAT ABSORBS
MAXIMALLY IN THE RED REGION OF THE SPECTRUM AND THE PFR FORM THAT
ABSORBS MAXIMALLY IN THE FAR-RED REGION. PHOTOCONVERSION OF PR IN
PFR INDUCES AN ARRAY OF MORPHOGENIC RESPONSES, WHEREAS
RECONVERSION OF PFR TO PR CANCELS THE INDUCTION OF THOSE
RESPONSES. PFR CONTROLS THE EXPRESSION OF A NUMBER OF NUCLEAR
GENES INCLUDING THOSE ENCODING THE SMALL SUBUNIT OF RIBULOSE-

```

```

CC BISPHOSPHATE CARBOXYLASE, CHLOROPHYLL A/B BINDING PROTEIN,
CC PROTOCHLOROPHYLLIDE REDUCTASE, RNA, ETC. IT ALSO CONTROLS
CC THE EXPRESSION OF ITS OWN GENE(S) IN A NEGATIVE FEEDBACK FASHION.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- PFM: CONTAINS ONE COVALENTLY LINKED TETRAPYRROLE CHROMOPHORE.
CC -1- SIMILARITY: BELONGS TO THE PHYTOCHROME FAMILY.
CC -1- SIMILARITY: CONTAINS 2 PAS (PER-ARNT-SIM) DIMERIZATION DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; AB018442; BAA74448.1; -
CC ENBL; U61207; ABA41996.1; -
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR003018; GAF.
DR InterPro; IPR004359; HIS_KIN_sig.
DR InterPro; IPR003661; His_kinA.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS_domain.
DR InterPro; IPR001294; Phytochrome.
DR Pfam; PF00360; phytochrome; 1.
DR Pfam; PF00512; signal; 1.
DR Pfam; PF00989; PAS; 2.
DR Pfam; PF01590; GAF; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR PRINTS; PR01033; PHYTOCHROME.
DR SMART; SM00065; GAF; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; HSKA; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR TIGRPFAM; TIGR00229; sensory_box; 2.
DR PROSITE; PS0109; HIS_KIN; 1.
DR PROSITE; PS0112; PAS; 2.
DR PROSITE; PS00245; PHYTOCHROME_1; 1.
DR PROSITE; PS0046; PHYTOCHROME_2; 1.
KW Transcription regulation; Photoreceptor; Phytochrome; Chromophore;
KW Repeat; Multigene family.
FT DOMAIN 620 690 PAS 1.
FT DOMAIN 750 824 PAS 2.
FT DOMAIN 904 1124 HISTIDINE KINASE.
FT BINDING 322 322 CHROMOPHORE (BY SIMILARITY).
FT CONFLICT 279 279 F -> S (IN REF. 2).
FT CONFLICT 292 292 C -> S (IN REF. 2).
SQ SEQUENCE 1137 AA; 125982 MW; P2A520181CFE7B32 CRC64;

Query Match 54.5%; Score 36; DB 1; Length 1137;
Best Local Similarity 36.4%; Pred. No. 54;
Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 11
DB 777 CLEWNEAMQKI 787

RESULT 15
ATX1_ARATH
ID ATX1_ARATH STANDARD; PRT; 1179 AA.
AC Q9LJ02.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Potential cation-transporting ATPase (EC 3.6.3.-).
GN AT5G23630 OR MQM1.11.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

C Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
C eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
X NCBI\_TaxID=3702;  
N [1]  
P SEQUENCE FROM N.A.  
C STRAIN=cv. Columbia;  
X MEDLINE=20181125; PubMed=10718197;  
A Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,  
A Tabata S.;  
T "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence  
T features of the regions of 3,076,755 bp covered by sixty P1 and TAC  
T clones";  
L DNA Res. 7:31-63(2000).  
C -!- FUNCTION: INVOLVED IN TRANSPORT OF CATIONS (POTENTIAL).  
C -!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.  
C -!- SUBCELLULAR LOCATION: Integral membrane protein.  
C -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY  
C (E1-E2 ATPASES). SUBFAMILY V.  
C -----  
C This SWISS-PROT entry is copyright. It is produced through a collaboration  
C between the Swiss Institute of Bioinformatics and the EMBL outstation -  
C the European Bioinformatics Institute. There are no restrictions on its  
C use by non-profit institutions as long as its content is in no way  
C modified and this statement is not removed. Usage by and for commercial  
C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
C or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
C -----  
C EMBL; AB025633; BAA97238.1; -  
C InterPro; IPR001757; ATPase\_E1-E2.  
C InterPro; IPR001454; Hlgase/hydriase.  
C Pfam; PF00122; E1-E2 ATPase; 1.  
C Pfam; PF00702; Hydrolase; 1.  
C PRINTS; PR00119; CATAPASE.  
C PROSITE; PS00154; ATPase\_E1\_E2; 1.  
W Hydrolase; transmembrane; Phosphorylation; ATP-binding; Metal-binding;  
W Magnesium.  
T DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).  
T TRANSNEM 21 42 POTENTIAL.  
T DOMAIN 43 50 EXTRACELLULAR (POTENTIAL).  
T TRANSNEM 51 71 POTENTIAL.  
T DOMAIN 72 192 CYTOPLASMIC (POTENTIAL).  
T TRANSNEM 193 215 POTENTIAL.  
T DOMAIN 216 218 EXTRACELLULAR (POTENTIAL).  
T TRANSNEM 219 238 POTENTIAL.  
T DOMAIN 239 402 CYTOPLASMIC (POTENTIAL).  
T TRANSNEM 403 422 POTENTIAL.  
T DOMAIN 423 435 EXTRACELLULAR (POTENTIAL).  
T TRANSNEM 436 453 POTENTIAL.  
T DOMAIN 454 947 CYTOPLASMIC (POTENTIAL).  
T TRANSNEM 948 967 POTENTIAL.  
T DOMAIN 968 979 EXTRACELLULAR (POTENTIAL).  
T TRANSNEM 980 997 POTENTIAL.  
T DOMAIN 998 1013 CYTOPLASMIC (POTENTIAL).  
T TRANSNEM 1014 1034 POTENTIAL.  
T DOMAIN 1035 1059 EXTRACELLULAR (POTENTIAL).  
T TRANSNEM 1060 1079 POTENTIAL.  
T DOMAIN 1080 1092 CYTOPLASMIC (POTENTIAL).  
T TRANSNEM 1093 1110 POTENTIAL.  
T DOMAIN 1111 1128 EXTRACELLULAR (POTENTIAL).  
T TRANSNEM 1129 1148 POTENTIAL.  
T DOMAIN 1149 1178 CYTOPLASMIC (POTENTIAL).  
T TRANSNEM 1175 1178 POLY-LYS.  
T MOD\_RES 491 491 PHOSPHORYLATION (BY SIMILARITY).  
T METAL 812 812 MAGNESIUM (BY SIMILARITY).  
T METAL 816 816 MAGNESIUM (BY SIMILARITY).  
Q SEQUENCE 1179 AA; 131115 MW; 4A3E82D222A4D78 CRC64;  
Query Match 54.5%; Score 36; DB 1; Length 1179;  
Best Local Similarity 44.4%; Pred. No. 56;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
Y 1 CFQWQXR 9  
|:|:|:

Db 1144 CYSWERLLR 1152

Search completed: February 21, 2003, 07:51:37  
Job time : 5.2093 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:44:43 ; Search time 21.6744 Seconds  
(without alignments)  
114.078 Million cell updates/sec

Title: US-09-743-107b-90  
Perfect score: 66  
Sequence: 1 CFQWQXMRKV 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_invertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacterioph.\*
- 17: sp\_archaea.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	89.4	711	4 Q8TCD2	Q8tcd2 homo sapien
2	55	83.3	38	4 Q9UCY5	Q9ucy5 homo sapien
3	50	75.8	33	6 Q9TR80	Q9tr80 ovis aries
4	41	62.1	105	10 Q9XFD5	Q9xfid5 cryza sativ
5	40	60.6	148	10 Q9XHP1	Q9xhp1 sesamum ind
6	40	60.6	275	5 Q93780	Q93780 caenorhabdi
7	40	60.6	932	5 Q19153	Q19153 caenorhabdi
8	39	59.1	203	16 Q986A0	Q986a0 rhizobium l
9	39	59.1	206	16 Q8UHC2	Q8uhc2 agrobacteri
10	39	59.1	206	16 Q92RH8	Q92rh8 rhizobium m
11	39	59.1	208	16 Q8VFK3	Q8vfk3 bruceella me
12	39	59.1	279	16 Q8XSE2	Q8xse2 ralstonia s
13	39	59.1	306	4 Q8TAX2	Q8tax2 homo sapien
14	39	59.1	466	4 Q9NUS2	Q9nus2 homo sapien
15	38	57.6	91	15 Q77855	Q77855 human immun
16	38	57.6	91	15 Q77856	Q77856 human immun

Q96m2i homo sapien  
Q96cn2 arabidopsis  
Q22185 arabidopsis  
Q22188 arabidopsis  
Q9sws6 lycopersico  
Q8z536 salmonella  
Q9yqcl human immun  
Q9yqco human immun  
Q9yqb9 human immun  
Q9yqb8 human immun  
Q9yji7 human immun  
Q9yji2 human immun  
Q9g910 human immun  
Q8q454 human immun  
Q8uek3 agrobacteri  
Q9ate4 caulobacteri  
Q44473 caenorhabdi  
Q9nzw0 homo sapien  
Q9nzw3 homo sapien  
P96223 mycobacteri  
Q9lha1 arabidopsis  
Q9hpa3 halobacteri  
Q94937 homo sapien  
Q62582 encephalito  
Q8srg3 encephalito  
Q8sqi6 encephalito  
Q9fhi9 arabidopsis  
Q9imq5 lumpy skin  
Q9atcp0 oryza sativ

ALIGNMENTS

RESULT 1  
Q8TCD2 ID Q8TCD2 PRELIMINARY; PRT; 711 AA.  
AC Q8TCD2;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Lactotraneferri.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN NCBI\_TaxID=9606;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PROSTATE;  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC022347; AAH2347.1; -;  
SQ SEQUENCE 711 AA; 78327 MW; 1B9C7EB097C45PAF CRC64;

Query Match 89.4%; Score 59; DB 4; Length 711;  
Best Local Similarity 90.9%; Pred. No. 0.0067; 1; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 1;

Qy 1 CFQWQXMRKV 11  
Db 39 CFQWQXMRKV 49

RESULT 2  
Q9UCY5 ID Q9UCY5 PRELIMINARY; PRT; 38 AA.  
AC Q9UCY5;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE Lactoferrin homolog (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OK NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96081613; PubMed=8551695;  
 RY Sato I.;  
 RT "Characterization of the 84-kDa protein with ABH activity in human  
 seminal plasma";  
 RL Jpn. J. Legal Med. 49:281-293 (1995).  
 DR HSP; P02788; IBKA.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 1.  
 SQ SEQUENCE 38 AA; 4459 MW; 0402F490B5EBDEB CRC64;  
 Query Match 83.3%; Score 55; DB 4; Length 38;  
 Best Local Similarity 50.9%; Pred. No. 0.0019;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 2 FQWQKMKVR 12  
 Db 21 FQWQKMKVR 31  
 RESULT 3  
 Q9TR80 PRELIMINARY; PRT; 33 AA.  
 ID Q9TR80;  
 AC Q9TR80;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DE Lactoferrin (Fragment).  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OK NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=95127729; PubMed=7827104;  
 RA Qian Z.Y., Jolles P., Migliore-Samouir D., Fiat A.M.;  
 RL Biochim. Biophys. Acta 1243:25-32 (1995).  
 DR HSP; O77698; ICE2  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 1.  
 SQ SEQUENCE 33 AA; 3914 MW; D1904CAE15A73961 CRC64;  
 Query Match 75.8%; Score 50; DB 6; Length 33;  
 Best Local Similarity 63.6%; Pred. No. 0.014;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 CFQWQKMKRV 11  
 Db 19 CYWQKMKRL 29  
 RESULT 4  
 Q9XFD5 PRELIMINARY; PRT; 105 AA.  
 ID Q9XFD5;  
 AC Q9XFD5;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DE Cytochrome p450 (Fragment).  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzae; Oryza.  
 OK NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PANICLE;  
 RA Liu J., Yang J.;  
 RT "Suppressor subtractive hybridization (SSH) identified candidate

RT genes that are differentially expressed at rice young panicle.";  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.

DR EMBL; AF140486; AAD29699.1; -.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR Pfam; PF00667; P450; 1.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME P450; UNKNOWN\_1.  
 KW Heme; Monooxygenase; Oxidoreductase.  
 FT NON\_TER 1  
 SQ SEQUENCE 105 AA; 11912 MW; B0BEFCD487E19F9 CRC64;

Query Match 62.1%; Score 41; DB 10; Length 105;  
 Best Local Similarity 60.0%; Pred. No. 2.3;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQWQKMKR 10  
 Db 61 CFQWRLGKK 70

## RESULT 5

Q9XHP1 PRELIMINARY; PRT; 148 AA.  
 ID Q9XHP1;  
 AC Q9XHP1;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE 2S albumin.  
 OS Sesamum indicum (Oriental sesame) (Gingelly).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Lamiales; Pedaliaceae; Sesamum.  
 OK NCBI\_TaxID=4182;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TAINAN 1;  
 RX MEDLINE=20074970; PubMed=10606554;  
 RA Tai S.S.K., Wu L.S.H., Chen E.C.F., Tzen J.T.C.;  
 RT "Molecular cloning of 11S globulin and 2S albumin, the two major seed  
 storage proteins in sesame";  
 RL J. Agric. Food Chem. 47:4932-4938 (1999).  
 DR EMBL; AF091841; AAD42943.1; -.  
 DR InterPro; IPR003612; AAI.  
 DR InterPro; IPR000617; Napin.  
 DR InterPro; IPR001768; Try/amy1 inhbr.  
 DR Pfam; PF00234; try/alpha\_aml; 1.  
 DR PRINTS; PR00496; NAEIN.  
 DR SMART; SM00499; AAI; 1.  
 SQ SEQUENCE 148 AA; 17524 MW; BA46B033BA13E3DE CRC64;

Query Match 60.6%; Score 40; DB 10; Length 148;  
 Best Local Similarity 56.7%; Pred. No. 5.1;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CFQWQKMKR 9  
 Db 54 CMQWRSR 62

## RESULT 6

Q93780 PRELIMINARY; PRT; 275 AA.  
 ID Q93780;  
 AC Q93780;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE P53H4.4 protein.  
 GN P53H4.4  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Telodermata; Caenorhabditis.  
 OK NCBI\_TaxID=6239;



```

IN SEQUENCE FROM N.A.
IP Dobson R.;
IL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
IN [2]
IP SEQUENCE FROM N.A.
IX MEDLINE=99069613; PubMed=9851916;
IA none;
IT "Genome sequence of the nematode C.elegans: A platform for
IT investigating biology.";
IL Science 282:2012-2018(1998).
IX EMBL; Z81089; CAB03137.1; -.
IQ SEQUENCE 275 AA; 31716 MW; 9D69524FFC704DB2 CRC64;

Query Match 60.6%; Score 40; DB 5; Length 275;
Best Local Similarity 63.6%; Pred. No. 9; 7;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Y 2 FQWQXMRKVR 12
|||: |||
b 262 FQWISMRKTR 272

RESULT 7
I19153 PRELIMINARY; PRT; 932 AA.
C Q19153
T 01-NOV-1996 (TrEMBLrel. 01, Created)
T 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
T 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
E Hypothetical 105.1 kDa protein.
F07C3.1.
S Caenorhabditis elegans.
C Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;
C Rhabditidae; Peloderinae; Caenorhabditis.
X NCBI_TaxID=6239;
N [1]
P SEQUENCE FROM N.A.
C STRAIN=BRISTOL N2;
X MEDLINE=99069613; PubMed=9851916;
A None;
T "Genome sequence of the nematode C. elegans: a platform for
T investigating biology. The C. elegans Sequencing Consortium.";
L Science 282:2012-2018(1998).
N [2]
P SEQUENCE FROM N.A.
C STRAIN=BRISTOL N2;
C Faveillo A., Gattung S.;
T "The sequence of C. elegans cosmid F07C3.";
L Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
N [3]
P SEQUENCE FROM N.A.
C STRAIN=BRISTOL N2;
A Waterston R.;
T "Direct Submission.";
L Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
R EMBL; U50308; AAG24025.1; -.
R InterPro; IPR000731; HMGCR/patch_5TM.
R PROSITE; PS50156; SSD; 1.
W Hypothetical protein.
Q SEQUENCE 932 AA; 105144 MW; 66680619ADACBFD5 CRC64;

Query Match 60.6%; Score 40; DB 5; Length 932;
Best Local Similarity 63.6%; Pred. No. 34;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Y 2 FQWQXMRKVR 12
|||: |||
b 579 FQWQSRARLVK 589

RESULT 8
986A0

```

```

ID Q986A0 PRELIMINARY; PRT; 205 AA.
AC Q986A0;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Pyridoxamine 5'-phosphate oxidase.
GN MLL7454.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuro A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003011; BAB53553.1; -.
DR InterPro; IPR000659; Pyridox oxidase.
DR Pfam; PF01243; Pyridox oxidase; 1.
DR ProDom; PD006312; Pyridox oxidase; 1.
DR TIGRFAMs; TIGR00558; pdsH_1.
DR PROSITE; PS01064; PYRIDOX_OXIDASE; 1.
KW Complete proteome.
SQ SEQUENCE 205 AA; 23300 MW; 0BADE4CD312327EA CRC64;

Query Match 59.1%; Score 39; DB 16; Length 205;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CFQWQXMRKVR 12
|||: |||
Db 88 CFHWKSLRQVR 99

RESULT 9
Q8UHC2 PRELIMINARY; PRT; 206 AA.
ID Q8UHC2
AC Q8UHC2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Pyridoxamine 5'-phosphate oxidase.
GN PDXH OR ATU0760 OR AGR_C_1381.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eissen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley P., Tingey S.V., Tomb J.-F., Gordon M.P., Olsen M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;

```

RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
 RA Quorillo B., Gordon B.S., Cao Y., Askenazi M., Halling C., Mullin L.,  
 RA Houmlel K., Goldman J., Vaudin M., Iatchouk O., Epp A., Liu F.,  
 RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,  
 RA Flanagan C., Crowell C., Garson J., Lomo C., Sear C., Strub G.,  
 RA Cielo C., Slater S.,  
 RT "Genome sequence of the plant pathogen and biotechnology agent  
 RT Agrobacterium tumefaciens C58,"  
 RL Science 294:2323-2328(2001).  
 DR EMBL: AE009043; AAL41776.1; -;  
 DR EMBL: AE008009; AAK85569.1; -;  
 KW Complete proteome.  
 SQ SEQUENCE 206 AA; 23720 MW; 3BE48BAE5307C0C1 CRC64;

Query Match 59.1%; Score 39; DB 16; Length 206;  
 Best Local Similarity 50.0%; Pred. No. 11;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWQXMRKVR 12  
 |||:|:|:  
 Db 88 CFHWKSLRRQVR 99

RESULT 10  
 Q92RH8 PRELIMINARY; PRT; 206 AA.  
 AC Q92RH8;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Probable pyridoxamine 5'-phosphate oxidase (PNP/PMF oxidase) protein  
 DE (EC 1.4.3.5).  
 DE PDHX OR R00895 OR SMC00069.  
 GN Rhizobium meliloti (Sinorhizobium meliloti).  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Sinorhizobium.  
 OX NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021;  
 RX MEDLINE=21396507; PubMed=11481430;  
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,  
 RA Pohl T., Portetlelle D., Puehler A., Purnelle B., Ramsperger U.,  
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.,  
 RT "Analysis of the chromosome sequence of the legume symbiont  
 RT Sinorhizobium meliloti strain 1021,"  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 DR EMBL: AL591785; CAC45467.1; -;  
 DR InterPro; IPR000659; Pyridox oxidase.  
 DR Pfam; PF01243; Pyridox oxidase; 1.  
 DR ProDom; PD006312; Pyridox oxidase; 1.  
 DR TIGRFAMs; TIGR00558; pdxH; 1.  
 DR PROSITE; PS01064; PYRIDOX\_OXIDASE; UNKNOWN 1.  
 KW Oxidoreductase; Complete proteome.  
 SQ SEQUENCE 206 AA; 23900 MW; A2DB74229DACA97A CRC64;

Query Match 59.1%; Score 39; DB 16; Length 206;  
 Best Local Similarity 50.0%; Pred. No. 11;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWQXMRKVR 12  
 |||:|:|:  
 Db 88 CFHWKSLRRQVR 99

RESULT 11  
 Q8YFK3 PRELIMINARY; PRT; 208 AA.  
 AC Q8YFK3;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Probable pyridoxamine 5'-phosphate oxidase (EC 1.4.3.5).  
 GN BME11517.  
 OS Brucella melitensis  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Brucellaceae; Brucella.  
 OX NCBI\_TaxID=29459;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=16M / ATCC 23456 / BIOTYPE 1;  
 RX MEDLINE=20020109; PubMed=11756688;  
 RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Muijer C., Los T.,  
 RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,  
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,  
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,  
 RA Hasselkorn R., Kyprides N., Overbeek R.,  
 RT "The genome sequence of the facultative intracellular pathogen  
 RT Brucella melitensis,"  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).  
 DR EMBL: AE009587; AAL52698.1; -;  
 DR InterPro; IPR000659; Pyridox oxidase.  
 DR Pfam; PF01243; Pyridox oxidase; 1.  
 DR ProDom; PD006312; Pyridox oxidase; 1.  
 DR TIGRFAMs; TIGR00558; pdxH; 1.  
 DR PROSITE; PS01064; PYRIDOX\_OXIDASE; 1.  
 KW Oxidoreductase; Complete proteome.  
 SQ SEQUENCE 208 AA; 23866 MW; CBIF50BC9612DE28 CRC64;

Query Match 59.1%; Score 39; DB 16; Length 208;  
 Best Local Similarity 50.0%; Pred. No. 11;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWQXMRKVR 12  
 |||:|:|:  
 Db 90 CFHWKSLRRQVR 101

RESULT 12  
 Q8XSE2 PRELIMINARY; PRT; 279 AA.  
 AC Q8XSE2;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Putative ICC protein homolog.  
 GN ICC OR RS0534 OR RS00414.  
 OS Ralstonia solanacearum (Pseudomonas solanacearum).  
 OC Plasmid megaplasmid.  
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;  
 OC Ralstonia.  
 OX NCBI\_TaxID=305;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GM11000;  
 RX MEDLINE=21681879; PubMed=11823852;  
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
 RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,  
 RA Chandler C., Choisme N., Claudel-Renard C., Cunha S., Demange N.,  
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,  
 RA Siguer P., Thebaud P., Whalen M., Wincker P., Levy M.,  
 RA Weissenbach J., Boucher C.A.,  
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum,"  
 RL Nature 415:497-502(2002).  
 DR EMBL: AL646079; CAD17685.1; -;  
 DR InterPro; IPR004843; M-ppetrase.  
 DR InterPro; IPR004844; S/T\_phosphatase.  
 DR Pfam; PF00149; Metallophos; 1.  
 KW Plasmid; Complete proteome.  
 SQ SEQUENCE 279 AA; 31541 MW; ABB39818004B2EDA CRC64;

Query Match 59.1%; Score 39; DB 16; Length 279;  
 Best Local Similarity 41.7%; Pred. No. 15;  
 Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Query Match 59.1%; Score 39; DB 4; Length 466;  
 Best Local Similarity 55.6%; Pred. No. 26;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQXMR 9  
 |||||:|  
 DB 269 CFQWESTLR 277

## RESULT 15

Q77855 PRELIMINARY; PRT; 91 AA.  
 AC Q77855;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Envelope protein (Fragment).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95191002; PubMed=7884875;  
 RA Mulder-Kampinga G.A., Simonon A., Kuiken C.L., Dekker J.,  
 RA Scherpbier H.J., de Perre P., Boer K., Goudsmit J.;  
 RT "Similarity in env and gag genes between genomic RNAs of human  
 RT immunodeficiency virus type 1 (HIV-1) from mother and infant is  
 RT unrelated to time of HIV-1 RNA positivity in the child.";  
 RL J. Virol. 69:2285-2296(1995).  
 DR EMBL; Z47867; CAA87881.1; -.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 KW AIDS; Coat protein; Glycoprotein.  
 FT NON TER 1 1  
 FT NON TER 91 91  
 SQ SEQUENCE 91 AA; 10530 MW; 8B10C62011F305D6 CRC64;

Query Match 57.6%; Score 38; DB 15; Length 91;  
 Best Local Similarity 60.0%; Pred. No. 7.4;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 QWQXMRKVR 12  
 |||||:|  
 DB 67 QWNETLQKVR 76

Search completed: February 21, 2003, 08:00:44  
 Job time : 22.6744 secs

QY 1 CFQWQXMRKVR 12  
 |||||:|  
 DB 244 CFQWEXGTRTAK 255

## RESULT 13

Q8TAX2 PRELIMINARY; PRT; 306 AA.  
 AC Q8TAX2;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DE Similar to hypothetical protein FLJ11175.  
 OS Homo sapiens (Human).  
 DC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 DC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 DX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PANCREAS;  
 RA Strausberg R.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC025708; AAH25708.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 306 AA; 34551 MW; 43D9AD722566B343 CRC64;

Query Match 59.1%; Score 39; DB 4; Length 306;  
 Best Local Similarity 55.6%; Pred. No. 17;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQXMR 9  
 |||||:|  
 DB 269 CFQWESTLR 277

## RESULT 14

Q9NUS2 PRELIMINARY; PRT; 466 AA.  
 AC Q9NUS2;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DE CDNA FLJ11175 fis, clone PLACE1007375, weakly similar to phorbol  
 DE ester/diacylglycerol-binding protein UNC-13.  
 OS Homo sapiens (Human).  
 DC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 DC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 DX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
 RA Watanabe M., Hosoi T., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
 RA Ninomiya K., Iwayanagi T.;  
 RL "NEDO human cDNA sequencing project.";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK002037; BAA92048.1; -.  
 DR HSSP; P21707; 1BYN.  
 DR InterPro; IPR000008; C2.  
 DR InterPro; IPR000504; RNA\_rec\_mot.  
 DR Pfam; PF00168; C2; 1.  
 DR PRINTS; PR00360; C2DOMAIN.  
 DR SMART; SM00239; C2; 1.  
 DR PROSITE; PS00499; C2\_DOMAIN\_1; UNKNOWN\_1.  
 DR PROSITE; PS50004; C2\_DOMAIN\_2; 1.  
 DR PROSITE; PS00030; RRM\_RNP\_1; UNKNOWN\_1.  
 QI SEQUENCE 466 AA; 53192 MW; E41133A5062F58D6E CRC64;

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: February 21, 2003, 07:37:21 ; Search time 28.093 Seconds  
(without alignments)  
56.918 Million cell updates/sec

Title: US-09-743-107B-91

Perfect score: 71

Sequence: 1 CFQWKNWRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002.\*  
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	71	100.0	12	21	AAV78091 Human lactoferrin
2	71	100.0	12	21	AAV78092 Human lactoferrin
3	67	94.4	12	21	AAV78038 Human lactoferrin
4	67	94.4	12	21	AAV78046 Human lactoferrin
5	67	94.4	12	21	AAV78047 Human lactoferrin
6	67	94.4	12	21	AAV78084 Human lactoferrin
7	67	94.4	13	21	AAV78037 Human lactoferrin
8	67	94.4	13	21	AAV78048 Human lactoferrin
9	67	94.4	13	21	AAV78049 Human lactoferrin
10	67	94.4	14	21	AAV78036 Human lactoferrin

11	67	94.4	14	21	AAV78050 Human lactoferrin
12	67	94.4	14	21	AAV78051 Human lactoferrin
13	67	94.4	15	17	AAV78054 Peptide for anti-u
14	67	94.4	15	21	AAV78035 Human lactoferrin
15	67	94.4	15	21	AAV78062 Human lactoferrin
16	67	94.4	15	21	AAV78063 Human lactoferrin
17	67	94.4	16	21	AAV78031 Human lactoferrin
18	67	94.4	16	21	AAV78064 Human lactoferrin
19	67	94.4	16	21	AAV78065 Human lactoferrin
20	67	94.4	17	21	AAV78034 Human lactoferrin
21	67	94.4	17	21	AAV78066 Human lactoferrin
22	67	94.4	17	21	AAV78067 Human lactoferrin
23	67	94.4	18	15	AAV78067 Human lactoferrin
24	67	94.4	18	17	AAV78067 Human lactoferrin
25	67	94.4	18	17	AAV78033 Advanced glycosyla
26	67	94.4	19	21	AAV78033 Human lactoferrin
27	67	94.4	19	21	AAV78032 Amino acid sequenc
28	67	94.4	20	13	AAV78032 Human lactoferrin
29	67	94.4	20	14	AAV78032 Anti microbial pep
30	67	94.4	20	15	AAV78032 Lactoferrin-relate
31	67	94.4	20	15	AAV78032 Lactoferrin-derive
32	67	94.4	20	15	AAV78032 Lactoferrin-derive
33	67	94.4	20	15	AAV78032 Lactoferrin-derive
34	67	94.4	20	16	AAV78032 Bovine lactoferrin
35	67	94.4	20	16	AAV78032 Anti-parasitic lac
36	67	94.4	20	16	AAV78032 Anti-parasitic lac
37	67	94.4	20	16	AAV78032 Peptide for anti-u
38	67	94.4	20	17	AAV78032 Lactoferrin-derive
39	67	94.4	20	17	AAV78032 Lactoferrin-derive
40	67	94.4	20	17	AAV78032 Lactoferrin-derive
41	67	94.4	20	17	AAV78032 Lactoferrin-derive
42	67	94.4	20	17	AAV78032 Lactoferrin-derive
43	67	94.4	20	17	AAV78032 Lactoferrin-derive
44	67	94.4	20	18	AAV78032 Lactoferrin-derive
45	67	94.4	20	18	AAV78032 Anti-parasitic pep

#### ALIGNMENTS

RESULT 1  
AAV78091  
ID AAV78091 standard; Peptide; 12 AA.  
XX AC AAV78091;  
XX DT 25-APR-2000 (first entry)  
XX DE Human lactoferrin derived peptide SEQ ID NO:91.

Human; lactoferrin; modification; infection; inflammation; tumour;  
food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
urinary tract infection; colitis; Candida infection; fungicidal;  
bactericidal; preservative.  
XX OS Homo sapiens.  
XX OS Synthetic.  
XX FN WC200001730-A1.  
XX PD 13-JAN-2000.  
XX PF 06-JUL-1999; 99WO-SE01230.  
XX PR 06-JUL-1998; 98SE-0002441.  
XX PR 17-JUL-1998; 98SE-0002562.  
XX PR 29-DEC-1998; 98SE-0004614.  
XX PA (ASCI-) A\* SCI INVEST AB.  
XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections, -  
 XX PT Inflammations and tumors and for use in infant formula food -  
 XX PS Claim 22; Page 38; 102pp; English.

XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;  
 Query Match 100.0%; Score 71; DB 21; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWKRNMRKVR 12  
 DB 1 CFQWKRNMRKVR 12  
 |||||

RESULT 2  
 AAY78092 ID AAY78092 standard; Peptide; 12 AA.  
 AC AAY78092;  
 XX 25-APR-2000 (first entry)  
 XX Human lactoferrin derived peptide SEQ ID NO:92.  
 DE Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200001730-A1.  
 DN 13-JAN-2000.  
 XX 06-JUL-1999; 99WO-SE01230.  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections, -  
 PT inflammations and tumors and for use in infant formula food -  
 XX PS Claim 22; Page 38; 102pp; English.  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;  
 Query Match 100.0%; Score 71; DB 21; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWKRNMRKVR 12  
 DB 1 CFQWKRNMRKVR 12  
 |||||

RESULT 3  
 AAY78038 ID AAY78038 standard; Peptide; 12 AA.  
 AC AAY78038;  
 XX 25-APR-2000 (first entry)  
 XX Human lactoferrin derived peptide SEQ ID NO:38.  
 DE Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200001730-A1.  
 DN 13-JAN-2000.  
 XX 06-JUL-1999; 99WO-SE01230.  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections, -  
 PT inflammations and tumors and for use in infant formula food -  
 XX PS Claim 12; Page 70; 102pp; English.  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.

XX Sequence 12 AA;

Query Match 94.4%; Score 67; DB 21; Length 12;  
Best Local Similarity 91.7%; Pred. No. 9.7e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

2y 1 CFQWKNRKVR 12  
| | | | | | | | | |  
Db 1 CFQWKNRKVR 12

RESULT 4

ID AAY78046 standard; Peptide; 12 AA.

AC AAY78046;

25-APR-2000 (first entry)

Human lactoferrin derived peptide SEQ ID NO:46.

Human; lactoferrin; modification; infection; inflammation; tumour;  
food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
urinary tract infection; colitis; Candida infection; fungicidal;  
bactericidal; preservative.

Homo sapiens.  
Synthetic.

WO200001730-A1.

13-JAN-2000.

06-JUL-1999; 99WO-SE01230.

06-JUL-1998; 98SE-0002441.

17-JUL-1998; 98SE-0002562.

29-DEC-1998; 98SE-0004614.

(ASCI-) A+ SCI INVEST AB.

Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

WPI; 2000-147388/13.

New peptides used for treatment and prevention of infections,  
inflammations and tumors and for use in infant formula food -

Claim 15; Page 35; 102pp; English.

CC AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.

XX Sequence 12 AA;  
| | | | | | | | | |

Query Match 94.4%; Score 67; DB 21; Length 12;  
Best Local Similarity 91.7%; Pred. No. 9.7e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWKNRKVR 12  
| | | | | | | | | |  
Db 1 CFQWKNRKVR 12

RESULT 5

ID AAY78047 standard; Peptide; 12 AA.

AC AAY78047;

25-APR-2000 (first entry)

Human lactoferrin derived peptide SEQ ID NO:47.

Human; lactoferrin; modification; infection; inflammation; tumour;  
food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
urinary tract infection; colitis; Candida infection; fungicidal;  
bactericidal; preservative.

Homo sapiens.  
Synthetic.

WO200001730-A1.

13-JAN-2000.

06-JUL-1999; 99WO-SE01230.

06-JUL-1998; 98SE-0002441.

17-JUL-1998; 98SE-0002562.

29-DEC-1998; 98SE-0004614.

(ASCI-) A+ SCI INVEST AB.

Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

WPI; 2000-147388/13.

New peptides used for treatment and prevention of infections,  
inflammations and tumors and for use in infant formula food -

Claim 18; Page 73; 102pp; English.

CC AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.

XX Sequence 12 AA;

Query Match 94.4%; Score 67; DB 21; Length 12;  
Best Local Similarity 91.7%; Pred. No. 9.7e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWKNRKR 12  
 DB 1 CFQWQNRKR 12

## RESULT 6

AAV78084  
 ID AAV78084 standard; Peptide; 12 AA.

AC AAV78084;

DT 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:84.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

OS Homo sapiens.

OS Synthetic.

PN WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

PR 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

PS Claim 22; Page 36; 102pp; English.

XX AAV78001 to AAV78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 12 AA;

Query Match 94.4%; Score 67; DB 21; Length 12;

Best Local Similarity 91.7%; Pred. No. 9.7e-05;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWKNRKR 12

DB 1 CFQWQNRKR 12

## RESULT 7

AAV78037

ID AAV78037 standard; Peptide; 13 AA.

XX AAV78037;

DT 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:37.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

OS Homo sapiens.

OS Synthetic.

PN WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

PR 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

PS Claim 12; Page 70; 102pp; English.

XX AAV78001 to AAV78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 13 AA;

Query Match 94.4%; Score 67; DB 21; Length 13;

Best Local Similarity 91.7%; Pred. No. 0.00011;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWKNRKR 12

DB 2 CFQWQNRKR 13

## RESULT 8

AAV78048

ID AAV78048 standard; Peptide; 13 AA.

XX AAV78048;

DT 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:48.

X Human; lactoferrin; modification; infection; inflammation; tumour;  
W food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
W urinary tract infection; colitis; Candida infection; fungicidal;  
W bactericidal; preservative.  
X  
S Homo sapiens.  
S Synthetic.  
N WO200001730-A1.  
X 13-JAN-2000.  
X  
F 06-JUL-1999; 99WO-SE01230.  
X  
R 06-JUL-1998; 98SE-0002441.  
R 17-JUL-1998; 98SE-0002562.  
R 29-DEC-1998; 98SE-0004614.  
X (ASCI-) A+ SCI INVEST AB.  
X  
M Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
X WPI; 2000-147388/13.  
X  
R New peptides used for treatment and prevention of infections,  
X inflammations and tumors and for use in infant formula food -  
X Claim 15; Page 74; 102pp; English.  
X  
X AAY78001 to AAY78100 represent peptides having sequences based on human  
X lactoferrin. The peptides are taken up in the intestine through  
X binding to specific lactoferrin receptors and are then transported  
X through the circulation. A medicinal product of the peptide or fragment  
X can be used for treating and/or prevention of infections (such as  
X urinary tract infections, colitis, and Candida infection on a mucosal  
X membrane), inflammations and/or tumours. The peptides can also be used  
X in food stuffs such as infant formula food. The peptides are also  
X fungicidal and bactericidal and may also be used as preservatives.  
X Even though native human lactoferrin have been shown to have desired  
X anti-inflammatory anti-infectious and anti-tumoural properties they  
X cannot be used clinically on a broad basis because of high production  
X costs. Therefore, provision of peptides based on lactoferrin would  
X enable them to be used for the same purposes as lactoferrin at lower  
X cost.  
X  
X Query Match 94.4%; Score 67; DB 21; Length 13;  
X Best Local Similarity 91.7%; Pred. No. 0.00011;  
X Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
X  
X 1 CFQWKRNNRKVR 12  
X 2 CFQWQRNNRKVR 13  
X  
X RESULT 9  
X AAY78049  
X ID AAY78049 standard; Peptide; 13 AA.  
X  
X AAY78049;  
X  
X 25-APR-2000 (first entry)  
X  
X Human lactoferrin derived peptide SEQ ID NO:49.  
X  
X Human; lactoferrin; modification; infection; inflammation; tumour;  
X food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
X urinary tract infection; colitis; Candida infection; fungicidal;  
X bactericidal; preservative.  
X  
X Homo sapiens.

OS Synthetic.  
XX WO200001730-A1.  
XX 13-JAN-2000.  
XX  
PF 06-JUL-1999; 99WO-SE01230.  
XX  
PR 06-JUL-1998; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
XX  
PA (ASCI-) A+ SCI INVEST AB.  
XX  
PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX WPI; 2000-147388/13.  
XX  
PT New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
XX Claim 18; Page 74; 102pp; English.  
XX  
XX AAY78001 to AAY78100 represent peptides having sequences based on human  
XX lactoferrin. The peptides are taken up in the intestine through  
XX binding to specific lactoferrin receptors and are then transported  
XX through the circulation. A medicinal product of the peptide or fragment  
XX can be used for treating and/or prevention of infections (such as  
XX urinary tract infections, colitis, and Candida infection on a mucosal  
XX membrane), inflammations and/or tumours. The peptides can also be used  
XX in food stuffs such as infant formula food. The peptides are also  
XX fungicidal and bactericidal and may also be used as preservatives.  
XX Even though native human lactoferrin have been shown to have desired  
XX anti-inflammatory anti-infectious and anti-tumoural properties they  
XX cannot be used clinically on a broad basis because of high production  
XX costs. Therefore, provision of peptides based on lactoferrin would  
XX enable them to be used for the same purposes as lactoferrin at lower  
XX cost.  
XX  
XX Query Match 94.4%; Score 67; DB 21; Length 13;  
XX Best Local Similarity 91.7%; Pred. No. 0.00011;  
XX Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 1 CFQWKRNNRKVR 12  
XX 2 CFQWQRNNRKVR 13  
XX  
XX RESULT 10  
XX AAY78036  
XX ID AAY78036 standard; Peptide; 14 AA.  
XX  
XX AAY78036;  
XX  
XX 25-APR-2000 (first entry)  
XX  
XX Human lactoferrin derived peptide SEQ ID NO:36.  
XX  
XX Human; lactoferrin; modification; infection; inflammation; tumour;  
XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
XX urinary tract infection; colitis; Candida infection; fungicidal;  
XX bactericidal; preservative.  
XX  
XX Homo sapiens.  
XX Synthetic.  
XX WO200001730-A1.  
XX 13-JAN-2000.  
XX  
XX 06-JUL-1999; 99WO-SE01230.



XX PR 06-JUL-1998; 98SE-0002441.  
 XX PR 17-JUL-1998; 98SE-0002562.  
 XX PR 29-DEC-1998; 98SE-0004614.  
 XX PA (ASCI-) A+ SCI INVEST AB.  
 XX XX  
 XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX DR WPI; 2000-147388/13.  
 XX XX  
 XX PT New peptides used for treatment and prevention of infections, -  
 XX PT inflammations and tumors and for use in infant formula food  
 XX XX  
 XX PS Claim 12; Page 69; 102pp; English.  
 XX XX  
 XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 XX CC lactoferrin. The peptides are taken up in the intestine through  
 XX CC binding to specific lactoferrin receptors and are then transported  
 XX CC through the circulation. A medicinal product of the peptide or fragment  
 XX CC can be used for treating and/or prevention of infections (such as  
 XX CC urinary tract infections, colitis, and Candida infection on a mucosal  
 XX CC membrane), inflammations and/or tumors. The peptides can also be used  
 XX CC in food stuffs such as infant formula food. The peptides are also  
 XX CC fungicidal and bactericidal and may also be used as preservatives.  
 XX CC Even though native human lactoferrin have been shown to have desired  
 XX CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 XX CC cannot be used clinically on a broad basis because of high production  
 XX CC costs. Therefore, provision of peptides based on lactoferrin would  
 XX CC enable them to be used for the same purposes as lactoferrin at lower  
 XX CC cost.  
 XX XX  
 XX SQ Sequence 14 AA;  
 XX  
 Query Match 94.4%; Score 67; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 0.00011;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CFQWQNNRKRVR 12  
 Db 3 CFQWQNNRKRVR 14  
 |||||  
 RESULT 11  
 AAY78050  
 ID AAY78050 standard; Peptide; 14 AA.  
 XX AC AAY78050;  
 XX DT 25-APR-2000 (first entry)  
 XX DE Human lactoferrin derived peptide SEQ ID NO:50.  
 XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
 XX KW bactericidal; preservative.  
 XX OS Homo sapiens.  
 XX OS Synthetic.  
 XX PN WO200001730-A1.  
 XX PD 13-JAN-2000.  
 XX PF 06-JUL-1999; 99WO-SE01230.  
 XX PR 06-JUL-1998; 98SE-0002441.  
 XX PR 17-JUL-1998; 98SE-0002562.  
 XX PR 29-DEC-1998; 98SE-0004614.  
 XX XX  
 XX PA (ASCI-) A+ SCI INVEST AB.  
 XX XX  
 XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX DR WPI; 2000-147388/13.  
 XX XX  
 XX PT New peptides used for treatment and prevention of infections, -  
 XX PT inflammations and tumors and for use in infant formula food  
 XX XX  
 XX PS Claim 12; Page 69; 102pp; English.  
 XX XX  
 XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 XX CC lactoferrin. The peptides are taken up in the intestine through  
 XX CC binding to specific lactoferrin receptors and are then transported  
 XX CC through the circulation. A medicinal product of the peptide or fragment  
 XX CC can be used for treating and/or prevention of infections (such as  
 XX CC urinary tract infections, colitis, and Candida infection on a mucosal  
 XX CC membrane), inflammations and/or tumors. The peptides can also be used  
 XX CC in food stuffs such as infant formula food. The peptides are also  
 XX CC fungicidal and bactericidal and may also be used as preservatives.  
 XX CC Even though native human lactoferrin have been shown to have desired  
 XX CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 XX CC cannot be used clinically on a broad basis because of high production  
 XX CC costs. Therefore, provision of peptides based on lactoferrin would  
 XX CC enable them to be used for the same purposes as lactoferrin at lower  
 XX CC cost.  
 XX XX  
 XX SQ Sequence 14 AA;  
 XX  
 Query Match 94.4%; Score 67; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 0.00011;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CFQWQNNRKRVR 12  
 Db 3 CFQWQNNRKRVR 14  
 |||||  
 RESULT 11  
 AAY78050  
 ID AAY78050 standard; Peptide; 14 AA.  
 XX AC AAY78050;  
 XX DT 25-APR-2000 (first entry)  
 XX DE Human lactoferrin derived peptide SEQ ID NO:50.  
 XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
 XX KW bactericidal; preservative.  
 XX OS Homo sapiens.  
 XX OS Synthetic.  
 XX PN WO200001730-A1.  
 XX PD 13-JAN-2000.  
 XX PF 06-JUL-1999; 99WO-SE01230.  
 XX PR 06-JUL-1998; 98SE-0002441.  
 XX PR 17-JUL-1998; 98SE-0002562.  
 XX PR 29-DEC-1998; 98SE-0004614.  
 XX XX  
 XX PA (ASCI-) A+ SCI INVEST AB.  
 XX XX

PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 XX PT New peptides used for treatment and prevention of infections, -  
 XX PT inflammations and tumors and for use in infant formula food  
 XX XX  
 XX PS Claim 15; Page 75; 102pp; English.  
 XX XX  
 XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 XX CC lactoferrin. The peptides are taken up in the intestine through  
 XX CC binding to specific lactoferrin receptors and are then transported  
 XX CC through the circulation. A medicinal product of the peptide or fragment  
 XX CC can be used for treating and/or prevention of infections (such as  
 XX CC urinary tract infections, colitis, and Candida infection on a mucosal  
 XX CC membrane), inflammations and/or tumors. The peptides can also be used  
 XX CC in food stuffs such as infant formula food. The peptides are also  
 XX CC fungicidal and bactericidal and may also be used as preservatives.  
 XX CC Even though native human lactoferrin have been shown to have desired  
 XX CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 XX CC cannot be used clinically on a broad basis because of high production  
 XX CC costs. Therefore, provision of peptides based on lactoferrin would  
 XX CC enable them to be used for the same purposes as lactoferrin at lower  
 XX CC cost.  
 XX XX  
 XX SQ Sequence 14 AA;  
 XX  
 Query Match 94.4%; Score 67; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 0.00011;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CFQWQNNRKRVR 12  
 Db 3 CFQWQNNRKRVR 14  
 |||||  
 RESULT 12  
 AAY78051  
 ID AAY78051 standard; Peptide; 14 AA.  
 XX AC AAY78051;  
 XX DT 25-APR-2000 (first entry)  
 XX DE Human lactoferrin derived peptide SEQ ID NO:51.  
 XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
 XX KW bactericidal; preservative.  
 XX OS Homo sapiens.  
 XX OS Synthetic.  
 XX PN WO200001730-A1.  
 XX PD 13-JAN-2000.  
 XX PF 06-JUL-1999; 99WO-SE01230.  
 XX PR 06-JUL-1998; 98SE-0002441.  
 XX PR 17-JUL-1998; 98SE-0002562.  
 XX PR 29-DEC-1998; 98SE-0004614.  
 XX XX  
 XX PA (ASCI-) A+ SCI INVEST AB.  
 XX XX  
 XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX DR WPI; 2000-147388/13.  
 XX XX  
 XX PT New peptides used for treatment and prevention of infections, -  
 XX PT inflammations and tumors and for use in infant formula food  
 XX XX

18 Claim 18; Page 75; 102pp; English.  
 19 AAY78001 to AAY78100 represent peptides having sequences based on human  
 20 lactoferrin. The peptides are taken up in the intestine through  
 21 binding to specific lactoferrin receptors and are then transported  
 22 through the circulation. A medicinal product of the peptide or fragment  
 23 can be used for treating and/or prevention of infections (such as  
 24 urinary tract infections, colitis, and Candida infection on a mucosal  
 25 membrane), inflammations and/or tumours. The peptides can also be used  
 26 in food stuffs such as infant formula food. The peptides are also  
 27 fungicidal and bactericidal and may also be used as preservatives.  
 28 Even though native human lactoferrin have been shown to have desired  
 29 anti-inflammatory anti-infectious and anti-tumoural properties they  
 30 cannot be used clinically on a broad basis because of high production  
 31 costs. Therefore, provision of peptides based on lactoferrin would  
 32 enable them to be used for the same purposes as lactoferrin at lower  
 33 cost.  
 34  
 35 Query Match 94.4%; Score 67; DB 21; Length 14;  
 36 Best Local Similarity 91.7%; Pred. No. 0.00011;  
 37 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 38  
 39 1 CFQWQRNMRKVR 12  
 40 ||||:|||||  
 41 3 CFQWQRNMRKVR 14  
 42  
 43 RESULT 13  
 44 AAY78054  
 45 AAY78054 standard; Peptide; 15 AA.  
 46  
 47 AAY78054;  
 48  
 49 12-NOV-1996 (first entry)  
 50 Peptide for anti-ulcer agent.  
 51 anti-ulcer agent; low toxicity; stable; heat-resistant.  
 52 Synthetic.  
 53 JP08143468-A.  
 54 04-JUN-1996.  
 55 17-NOV-1994; 94JP-0283869.  
 56 17-NOV-1994; 94JP-0283869.  
 57 (MORG) MORINAGA MILK IND CO LTD.  
 58 WPI; 1996-318857/32.  
 59 Anti-ulcer agent contg. peptide - has low toxicity, is  
 60 heat-resistant and water-soluble  
 61 Claim 1; Page 11; 11pp; Japanese.  
 62 AAY78054 are peptides used in an anti-ulcer agent. The agent is low  
 63 in toxicity, is heat-resistant and stable in aqueous soln.. It can be  
 64 administered orally and be produced in large amounts.  
 65  
 66 Query Match 94.4%; Score 67; DB 17; Length 15;  
 67 Best Local Similarity 91.7%; Pred. No. 0.00012;  
 68 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 69  
 70 1 CFQWQRNMRKVR 12  
 71 ||||:|||||  
 72 2 CFQWQRNMRKVR 13

RESULT 14  
 AAY78035  
 ID AAY78035 standard; Peptide; 15 AA.  
 XX AAY78035;  
 XX 25-APR-2000 (first entry)  
 DT Human lactoferrin derived peptide SEQ ID NO:35.  
 DE Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.  
 OS WO2000001730-A1.  
 PN 13-JAN-2000.  
 PD 06-JUL-1999; 99WO-SE01230.  
 PF 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 DR New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 12; Page 69; 102pp; English.  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 CC  
 SQ Sequence 15 AA;  
 Query Match 94.4%; Score 67; DB 21; Length 15;  
 Best Local Similarity 91.7%; Pred. No. 0.00012;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 CFQWQRNMRKVR 12  
 ||||:|||||  
 Db 4 CFQWQRNMRKVR 15  
 RESULT 15  
 AAY78062  
 ID AAY78062 standard; Peptide; 15 AA.  
 XX AAY78062;  
 AC

XX 25-APR-2000 (first entry)  
DT Human lactoferrin derived peptide SEQ ID NO:62.  
DE  
DE Human lactoferrin; modification; infection; inflammation; tumour;  
XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX WO200001730-A1.  
PN  
XX 13-JAN-2000.  
PD  
XX 06-JUL-1999; 99WO-SE01230.  
PF  
XX 06-JUL-1998; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
XX  
PA (ASCI-) A+ SCI INVEST AB.  
XX  
XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
PI  
XX WPI; 2000-147388/13.  
DR  
XX  
XX New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
XX  
XX Claim 15; Page 81; 102pp; English.  
PS  
XX  
CC AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.  
XX  
XX  
SQ Sequence 15 AA;  
Query Match 94.4%; Score 67; DB 21; Length 15;  
Best Local Similarity 91.7%; Pred. No. 0.00012;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CFQWKRNRKVR 12  
Db 4 CFQWQRNRKVR 15  
|||||  
|||||

Search completed: February 21, 2003, 07:56:44  
Job time : 29.093 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

DM protein - protein search, using sw model

Run on: February 21, 2003, 07:50:40 ; Search time 8.93023 Seconds  
(without alignments)  
39.537 Million cell updates/sec

Title: US-09-743-107B-91  
Perfect score: 71  
Sequence: 1 CFQWKNRKRVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PGTUS-COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	67	94.4	18	1	US-08-204-487-3
2	67	94.4	18	2	US-08-485-948-8
3	67	94.4	18	2	US-08-628-380-8
4	67	94.4	18	2	US-08-475-055-8
5	67	94.4	20	1	US-07-755-161A-3
6	67	94.4	20	1	US-07-831-174-3
7	67	94.4	20	1	US-08-204-487-1
8	67	94.4	20	1	US-08-256-771-24
9	67	94.4	20	1	US-08-256-771-25
10	67	94.4	20	1	US-08-381-984-24
11	67	94.4	20	1	US-08-381-984-25
12	67	94.4	22	4	US-09-508-734-4
13	67	94.4	24	4	US-09-508-734-6
14	67	94.4	25	1	US-07-755-161A-10
15	67	94.4	25	1	US-08-891-174-10
16	67	94.4	25	1	US-08-204-487-7
17	67	94.4	29	1	US-09-508-734-8
18	67	94.4	36	1	US-07-755-161A-8
19	67	94.4	36	1	US-07-891-174-8
20	67	94.4	36	1	US-08-256-771-30
21	67	94.4	36	1	US-08-381-984-29
22	67	94.4	47	2	US-08-454-182A-6
23	67	94.4	47	2	US-08-406-271-6
24	67	94.4	50	2	US-08-693-274A-7
25	67	94.4	52	4	US-09-017-043A-3
26	67	94.4	53	2	US-08-454-182A-5
27	67	94.4	53	2	US-08-406-271-5

28	67	94.4	54	2	US-08-464-182A-2
29	67	94.4	54	2	US-08-406-271-2
30	67	94.4	634	3	US-08-724-586-2
31	67	94.4	694	4	US-09-421-633-2
32	67	94.4	694	4	US-09-932-190-2
33	67	94.4	705	2	US-08-655-640-2
34	67	94.4	708	2	US-08-655-640-4
35	67	94.4	711	1	US-08-154-019-4
36	67	94.4	711	1	US-08-461-333-4
37	67	94.4	711	3	US-08-464-167-4
38	67	94.4	711	3	US-09-158-313-4
39	67	94.4	711	4	US-08-476-798-4
40	64	90.1	711	1	US-08-145-681-2
41	64	90.1	711	1	US-08-250-308-2
42	64	90.1	711	1	US-08-453-703-2
43	64	90.1	711	2	US-08-456-106-2
44	64	90.1	711	3	US-08-456-108-2
45	64	90.1	711	4	US-09-265-577-2

ALIGNMENTS

RESULT 1  
US-08-204-487-3  
; Sequence 3, Application US/08204487  
; Patent No. 5563425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAOKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGEKI  
; APPLICANT: DOSAKO, SHUN ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; INHIBITORS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESSEE: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: RJN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..18  
; OTHER INFORMATION: /note= "HUMAN LACTOFERRIN PEPTIDE"

OTHER INFORMATION: (20-37)"

US-08-204-487-3

Query Match 94.4%; Score 67; DB 1; Length 18;  
Best Local Similarity 91.7%; Pred. No. 9.9e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWKNRKVR 12  
| | | | | | | | | |  
Db 1 CFQWQNRKVR 12

## RESULT 2

US-08-485-948-8  
; Sequence 8, Application US/08485948  
; Patent No. 5855882  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION NUMBER: US/08/485,948  
FILING DATE: APRIL 7, 1995  
CLASSIFICATION: 436

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/488,217  
FILING DATE: JUNE 7, 1995  
APPLICATION NUMBER: 08/418,642  
FILING DATE: APRIL 7, 1995  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 947-1-008A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521

INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
DESCRIPTION: LF-C1, 8-25  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal

US-08-485-948-8

Query Match 94.4%; Score 67; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 9.9e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWKNRKVR 12  
| | | | | | | | | |  
Db 1 CFQWQNRKVR 12

## RESULT 3

US-08-628-380-8  
; Sequence 8, Application US/08628380  
; Patent No. 5891341  
; GENERAL INFORMATION:

; APPLICANT: LI, YONG MING  
; APPLICANT: VLASSARA, HELEN  
; APPLICANT: CERAMI, ANTHONY  
; TITLE OF INVENTION: AGE-MEDIATED INHIBITION OF ANTIBACTERIAL PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION NUMBER: US/08/628,380  
FILING DATE: April 4, 1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/418,642  
FILING DATE: APRIL 7, 1995  
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 947-1-008 CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521

INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
DESCRIPTION: LF-C1, 8-25  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal

US-08-628-380-8

Query Match 94.4%; Score 67; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 9.9e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWKNRKVR 12  
| | | | | | | | | |  
Db 1 CFQWQNRKVR 12

## RESULT 4

US-08-475-055-8  
; Sequence 8, Application US/08475055  
; Patent No. 5962245  
; GENERAL INFORMATION:

; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/475,055  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/485,948  
FILING DATE:  
APPLICATION NUMBER: 08/488,217  
FILING DATE: JUNE 7, 1995  
APPLICATION NUMBER: 08/418,642  
FILING DATE: APRIL 7, 1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 947-1-008A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
DESCRIPTION: LF-C1, 8-25  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
JS-08-475-055-8  
Query Match 94.4%; Score 67; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 9.9e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
2y 1 CFQWKNRKVR 12  
2b 1 CFQWQNNRKVR 12  
RESULT 5  
JS-07-755-161A-3  
Sequence 3, Application US/07755161A  
Patent No. 5304833  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: Antimicrobial Peptides and an  
TITLE OF INVENTION: Antimicrobial Agent  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
COMPUTER: IBM Compatible  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/755,161A  
FILING DATE: 19910905  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX: 202-371-8856  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 2  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 19"  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 19  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 2"  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-755-161A-3  
Query Match 94.4%; Score 67; DB 1; Length 20;

```
Best Local Similarity 91.7%; Pred. No. 0.00011; Indels 0; Gaps 0;
Matches 11; Conservative 1; Mismatches 0;

Qy 1 CFQWKNMRKVR 12
    |||||:|||||
Db 2 CFQWQNRKVR 13

RESULT 6
US-07-891-174-3
; Sequence 3, Application US/07891174
; Patent No. 5317084
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; TITLE OF INVENTION: Antimicrobial Agent
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/891,174
; FILING DATE: 29-MAY-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/755,161
; FILING DATE: 05-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site

; LOCATION: 2
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 19"
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 19
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 2"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-07-891-174-3

Query Match 94.4%; Score 67; DB 1; Length 20;
Best Local Similarity 91.7%; Pred No. 0.00011;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWKNMRKVR 12
    |||||:|||||
Db 2 CFQWQNRKVR 13

RESULT 7
US-08-204-487-1
; Sequence 1, Application US/08204487
; Patent No. 5565425
; GENERAL INFORMATION:
; APPLICANT: YAMAMOTO, NAOKI
; APPLICANT: NAKASHIMA, HIDEKI
; APPLICANT: MOSUCHI, WATARU
; APPLICANT: TANAKA, SHIGERU
; APPLICANT: DOSAKO, SHUN-ICHI
; APPLICANT: KAWASAKI, YOSHITIRO
; APPLICANT: UCHIDA, TOSHIKI
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION
; TITLE OF INVENTION: INHIBITORS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
; ADDRESS: THIBEAULT
; STREET: 53 STATE STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,487
; FILING DATE: 02-MAR-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, PAULA A.
; REGISTRATION NUMBER: 32,503
; REFERENCE/DOCKET NUMBER: EJN-019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
```

TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1-20  
OTHER INFORMATION: /note= "ANTIBACTERIAL PEPTIDE  
DERIVED FROM HUMAN LACTOFERRIN"  
US-08-204-487-1

Query Match 94.4%; Score 67; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 0.00011;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRNMRKVR 12  
| | | | |  
DB 2 CFQWQRNMRKVR 13

RESULT 8  
US-08-256-771-24  
; Sequence 24, Application US/08256771  
; Patent No. 5656591  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
; TITLE OF INVENTION: PRODUCTS THEREWITH  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,771  
FILING DATE: July 22, 1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:

INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "Cys residues are linked by  
OTHER INFORMATION: disulfide bond"

US-08-256-771-24

Query Match 94.4%; Score 67; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 0.00011;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRNMRKVR 12  
| | | | |  
DB 2 CFQWQRNMRKVR 13

RESULT 9  
US-08-256-771-25  
; Sequence 25, Application US/08256771  
; Patent No. 5656591  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
; TITLE OF INVENTION: PRODUCTS THEREWITH  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,771  
FILING DATE: July 22, 1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:

INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "Cys residues are protected to  
OTHER INFORMATION: prevent disulfide bond"  
US-08-256-771-25

Query Match 94.4%; Score 67; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 0.00011;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRNMRKVR 12  
| | | | |  
DB 2 CFQWQRNMRKVR 13

RESULT 10  
US-08-381-984-24



Sequence 24, Application US/08381984  
Patent No. 5804555  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIOXIDANT  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/381,984  
FILING DATE: April 11, 1995  
CLASSIFICATION: 252  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "cysteine residues at positions 2  
OTHER INFORMATION: and 19 are bonded by disulfide linkage"  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "the specified peptide as well as  
OTHER INFORMATION: peptides including the specified peptide as a fragment thereof"  
US-08-381-984-24  
Query Match 94.4%; Score 67; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 0.00011;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CFQWRNMRKVR 12  
DB 2 CFQWRNMRKVR 13  
RESULT 11  
US-06-381-984-25  
Sequence 25, Application US/08381984  
Patent No. 5804555  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIOXIDANT  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack

STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/381,984  
FILING DATE: April 11, 1995  
CLASSIFICATION: 252  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "the specified peptide as well as  
OTHER INFORMATION: peptides including the specified peptide as a fragment thereof"  
US-08-381-984-25  
Query Match 94.4%; Score 67; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 0.00011;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CFQWRNMRKVR 12  
DB 2 CFQWRNMRKVR 13  
RESULT 12  
US-09-508-734-4  
Sequence 4, Application US/09508734  
Patent No. 6423509  
GENERAL INFORMATION:  
APPLICANT: Samyang Genex Corporation  
TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and  
TITLE OF INVENTION: useful microorganism thereof  
FILE REFERENCE: PA/SYG/00139  
CURRENT APPLICATION NUMBER: US/09/508,734  
CURRENT FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: PCI/KR99/00373  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: KR1998-29351  
PRIOR FILING DATE: 1998-07-13  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Kopatentin 1.71  
SEQ ID NO 4

```
/ LENGTH: 22
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-508-734-4
Query Match          94.4%; Score 67; DB 4; Length 22;
Best Local Similarity 91.7%; Pred. No. 0.00012;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWKNRKVR 12
Db 2 CFQWKNRKVR 13

RESULT 13
US-09-508-734-6
Sequence 6, Application US/09508734
Patent No. 6423509
GENERAL INFORMATION:
APPLICANT: Samyang Genex Corporation
TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and
FILE REFERENCE: PA/SYG/00139
CURRENT APPLICATION NUMBER: US/09/508,734
CURRENT FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/KR99/00373
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: KR1998-29351
PRIOR FILING DATE: 1998-07-13
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Kopatentin 1.71
SEQ ID NO 6
LENGTH: 24
TYPE: PRT
ORGANISM: Homo sapiens
US-09-508-734-6

Query Match          94.4%; Score 67; DB 4; Length 24;
Best Local Similarity 91.7%; Pred. No. 0.00013;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWKNRKVR 12
Db 3 CFQWKNRKVR 14

RESULT 14
US-07-755-161A-10
Sequence 10, Application US/07755161A
Patent No. 5304633
GENERAL INFORMATION:
APPLICANT: Mamoru TOMITA et al.
TITLE OF INVENTION: Antimicrobial Peptides and an
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/755,161A
FILING DATE: 19910905
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX: 202-371-8856
TELEX:
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY: modified site
LOCATION: 4
IDENTIFICATION METHOD:
OTHER INFORMATION: /note= "thiol group of
OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with
OTHER INFORMATION: thiol group of Cys residue at location 21"
FEATURE:
NAME/KEY: modified site
LOCATION: 21
IDENTIFICATION METHOD:
OTHER INFORMATION: /note= "thiol group of
OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with
OTHER INFORMATION: thiol group of Cys residue at location 4"
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-07-755-161A-10

Query Match          94.4%; Score 67; DB 1; Length 25;
Best Local Similarity 91.7%; Pred. No. 0.00013;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWKNRKVR 12
Db 4 CFQWKNRKVR 15
```

```
RESULT 15
US-07-891-174-10
; Sequence 10 Application US/07891174
; Patent No. 5317084
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; TITLE OF INVENTION: Antimicrobial Agent
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/891,174
; FILING DATE: 29-MAY-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/755,161
; FILING DATE: 05-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 4
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 21"
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 21

; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 4"
; PUBLICATION INFORMATION:
; AUTHORS:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-07-891-174-10
; Query Match 94.4%; Score 57; DB 1; Length 25;
; Best Local Similarity 91.7%; Pred. No. 0.00013;
; Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
; Qy 1 CFQWKRNMRKVR 12
; Db 4 CFQWKRNMRKVR 15
; Search completed: February 21, 2003, 08:04:26
; Job time : 8.93023 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

XM protein - protein search, using sw model

run on: February 21, 2003, 08:08:15 ; Search time 6.88372 Seconds  
(without alignments)  
54.162 Million cell updates/sec

Title: US-09-743-107B-91

Perfect score: 71

Sequence: 1 CFQWKNRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 156504 seqs, 31069816 residues

Total number of hits satisfying chosen parameters: 156504

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/prodata/2/pubpaa/US08\_NEW PUB.pep.\*
- 2: /cgn2\_6/prodata/2/pubpaa/PCF\_NEW PUB.pep.\*
- 3: /cgn2\_6/prodata/2/pubpaa/US06\_NEW PUB.pep.\*
- 4: /cgn2\_6/prodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/prodata/2/pubpaa/US07\_NEW PUB.pep.\*
- 6: /cgn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/prodata/2/pubpaa/PCFUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/prodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/prodata/2/pubpaa/US09\_NEW PUB.pep.\*
- 10: /cgn2\_6/prodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/prodata/2/pubpaa/US10\_NEW PUB.pep.\*
- 12: /cgn2\_6/prodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/prodata/2/pubpaa/US60\_NEW PUB.pep.\*
- 14: /cgn2\_6/prodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	67	94.4	15	9	US-09-798-869-2
2	67	94.4	25	9	US-09-798-869-20
3	67	94.4	694	9	US-10-023-096-2
4	59	83.1	15	9	US-09-798-869-6
5	50	70.4	15	9	US-09-798-869-3
6	50	70.4	25	9	US-09-798-869-23
7	42	59.2	15	9	US-09-798-869-7
8	41	57.7	15	9	US-09-798-869-4
9	41	57.7	25	9	US-09-798-869-22
10	39	54.9	333	9	US-09-796-753-26
11	38	53.5	15	9	US-09-798-869-8
12	38	53.5	15	9	US-09-798-869-29
13	38	53.5	15	9	US-09-798-869-30
14	38	53.5	21	10	US-09-864-761-47985
15	38	53.5	489	9	US-09-868-320-2
16	37	52.1	77	10	US-09-864-761-41002
17	37	52.1	86	9	US-09-738-626-5715
18	37	52.1	184	10	US-09-925-301-1248
19	37	52.1	338	9	US-09-978-295A-119

```

20 37 52.1 338 9 US-09-978-697-119 Sequence 119, App
21 37 52.1 338 9 US-09-978-192A-119 Sequence 119, App
22 37 52.1 338 9 US-09-999-832A-119 Sequence 119, App
23 37 52.1 338 9 US-09-978-189-119 Sequence 119, App
24 37 52.1 553 9 US-09-796-753-14 Sequence 14, Appl
25 37 52.1 553 10 US-09-981-649A-6 Sequence 6, Appl
26 37 52.1 553 10 US-09-981-649A-24 Sequence 24, Appl
27 37 52.1 554 10 US-09-981-649A-30 Sequence 30, Appl
28 37 52.1 554 10 US-09-981-649A-32 Sequence 32, Appl
29 37 52.1 559 10 US-09-981-649A-28 Sequence 28, Appl
30 35 49.3 301 9 US-10-080-960-11 Sequence 11, Appl
31 35 49.3 302 10 US-09-948-078-2 Sequence 2, Appl
32 35 49.3 351 10 US-09-853-625B-16 Sequence 16, Appl
33 35 49.3 1013 9 US-10-028-072-38 Sequence 38, Appl
34 35 49.3 1013 9 US-10-121-043-38 Sequence 38, Appl
35 35 49.3 1013 9 US-10-123-904-38 Sequence 38, Appl
36 35 49.3 1013 9 US-10-140-470-38 Sequence 38, Appl
37 35 49.3 1013 9 US-10-175-746-38 Sequence 38, Appl
38 35 49.3 1013 9 US-10-176-918-38 Sequence 38, Appl
39 35 49.3 1013 9 US-10-176-921-38 Sequence 38, Appl
40 35 49.3 1013 9 US-10-137-865-38 Sequence 38, Appl
41 35 49.3 1013 9 US-10-140-474-38 Sequence 38, Appl
42 34 47.9 62 10 US-09-815-242-12129 Sequence 12129, A
43 34 47.9 62 10 US-09-815-242-13026 Sequence 13026, A
44 34 47.9 509 10 US-09-879-957-194 Sequence 194, App
45 34 47.9 846 9 US-10-051-409-4 Sequence 4, Appl

```

#### ALIGNMENTS

```

RESULT 1
US-09-798-869-2
; Sequence 2, Application US/09798869
; Publication No. US20030022821A1
; GENERAL INFORMATION:
; APPLICANT: JOHN SIGURD SVENDSEN
; APPLICANT: (YSTEIN REKDAL
; APPLICANT: BALDUR SVEINBJ (RNSON
; APPLICANT: LARS VORLAND
; TITLE OF INVENTION: BIOACTIVE PEPTIDES
; FILE REFERENCE: A34049.PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/798,869
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: PCT/GB99/02851
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: GB9818938.4
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 15
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-798-869-2

```

```

Query Match 94.4%; Score 67; DB 9; Length 15;
Best Local Similarity 91.7%; Pred. No. 5.1e-05;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CFQWKNRKVR 12
Db 3 CFQWKNRKVR 14

```

```

RESULT 2
US-09-798-869-20
; Sequence 20, Application US/09798869
; Publication No. US20030022821A1
; GENERAL INFORMATION:
; APPLICANT: JOHN SIGURD SVENDSEN
; APPLICANT: (YSTEIN REKDAL
; APPLICANT: BALDUR SVEINBJ (RNSON

```

APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/G899/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 20  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-798-869-20

Query Match 94.4%; Score 67; DB 9; Length 25;  
Best Local Similarity 91.7%; Pred. No. 8.2e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWKNNMKVR 12  
Db 3 CFQWQNNMKVR 14

RESULT 3  
US-10-023-096-2  
; Sequence 2, Application US/10023096  
; Patent No. US20020160941A1  
; GENERAL INFORMATION:  
; APPLICANT: Kruzel, Marian L.  
; APPLICANT: Kurecki, Tomasz  
; APPLICANT: Gollnick, Paul D.  
; APPLICANT: Doyle, Darrell J.  
; TITLE OF INVENTION: Cloning, Expression, and Uses of Human  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Jacobson, Price, Holman & Stern  
; STREET: 400 Seventh St. N.W.  
; CITY: Washington D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25.  
; CURRENT APPLICATION DATA:  
; FILING DATE: US/10/023,096

CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/724,586  
; FILING DATE: 30-SEPT-1996  
; APPLICATION NUMBER: US 08/238,445  
; FILING DATE: 05-MAY-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plaver, William E.  
; REGISTRATION NUMBER: 31,409  
; REFERENCE/DOCKET NUMBER: 10505/P58185C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 638-6666  
; TELEFAX: (202) 393-5350  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 694 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-10-023-096-2

Query Match 94.4%; Score 67; DB 9; Length 694;  
Best Local Similarity 91.7%; Pred. No. 0.0017;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWKNNMKVR 12  
Db 22 CFQWQNNMKVR 33

RESULT 4  
US-09-798-869-6  
; Sequence 6, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/G899/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; TYPE: PRT  
; LENGTH: 15  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of Homo sapiens  
; OTHER INFORMATION: sequence)  
US-09-798-869-6

Query Match 83.1%; Score 59; DB 9; Length 15;  
Best Local Similarity 83.3%; Pred. No. 0.001;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWKNNMKVR 12  
Db 3 CFQWQNNMKVR 14

RESULT 5  
US-09-798-869-3  
; Sequence 3, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/G899/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: CAPRINE  
US-09-798-869-3

Query Match 70.4%; Score 50; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.028;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWKNRKV 11  
|:|:|:|:  
DB 3 CYQWQRMKRL 13

RESULT 6  
US-09-798-869-23  
; Sequence 23, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: CAPRINE  
US-09-798-869-23

Query Match 70.4%; Score 50; DB 9; Length 25;  
Best Local Similarity 63.6%; Pred. No. 0.045;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWKNRKV 11  
|:|:|:|:  
DB 3 CYQWQRMKRL 13

RESULT 7  
US-09-798-869-7  
; Sequence 7, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURES:  
; OTHER INFORMATION: synthetic peptide (modified form of homo sapiens  
; OTHER INFORMATION: sequence)  
US-09-798-869-7

Query Match 59.2%; Score 42; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.055;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWKNRKV 11  
|:|:|:|:  
DB 3 CYQWQRMKRL 13

RESULT 8  
US-09-798-869-4  
; Sequence 4, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-798-869-4

Query Match 57.7%; Score 41; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.8;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWKNRKV 11  
|:|:|:|:  
DB 3 CLRQWNEKRV 13

RESULT 9  
US-09-798-869-22  
; Sequence 22, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 22  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-798-869-22

Query Match 57.7%; Score 41; DB 9; Length 25;  
Best Local Similarity 54.5%; Pred. No. 1.3;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWKNRKV 11  
|:|:|:|:  
DB 3 CLRQWNEKRV 13

## RESULT 10

US-09-796-753-26  
; Sequence 26, Application US/09796753  
; Publication No. US20030027998A1  
; GENERAL INFORMATION:  
; APPLICANT: McCarthy, Sean A.  
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF  
; FILE REFERENCE: 7853-227-999  
; CURRENT APPLICATION NUMBER: US/09/796,753  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 09/183,175  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: 09/223,094  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/223,546  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/224,246  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/259,388  
; PRIOR FILING DATE: 1999-02-26  
; PRIOR APPLICATION NUMBER: 60/122,458  
; PRIOR FILING DATE: 1999-03-01  
; PRIOR APPLICATION NUMBER: 09/312,359  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 09/336,536  
; PRIOR FILING DATE: 1999-06-18  
; PRIOR APPLICATION NUMBER: 09/342,687  
; PRIOR FILING DATE: 1999-06-29  
; PRIOR APPLICATION NUMBER: 09/345,464  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: 09/365,164  
; PRIOR FILING DATE: 1999-07-30  
; PRIOR APPLICATION NUMBER: 09/399,723  
; PRIOR FILING DATE: 1999-08-20  
; PRIOR APPLICATION NUMBER: 09/409,634  
; PRIOR FILING DATE: 1999-09-30  
; PRIOR APPLICATION NUMBER: 09/471,179  
; PRIOR FILING DATE: 1999-12-23  
; PRIOR APPLICATION NUMBER: 09/474,071  
; PRIOR FILING DATE: 1999-12-29  
; PRIOR APPLICATION NUMBER: 09/474,072  
; PRIOR FILING DATE: 1999-12-29  
; PRIOR APPLICATION NUMBER: 09/514,010  
; PRIOR FILING DATE: 2000-02-25  
; PRIOR APPLICATION NUMBER: 09/516,745  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 09/572,002  
; PRIOR FILING DATE: 2000-05-14  
; PRIOR APPLICATION NUMBER: 09/597,993  
; PRIOR FILING DATE: 2000-06-19  
; PRIOR APPLICATION NUMBER: 09/599,596  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 09/630,334  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: 09/606,565  
; PRIOR FILING DATE: 2000-06-29  
; PRIOR APPLICATION NUMBER: 09/606,317  
; PRIOR FILING DATE: 2000-06-29  
; PRIOR APPLICATION NUMBER: 09/665,665  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: 09/677,751  
; PRIOR FILING DATE: 2000-09-30  
; NUMBER OF SEQ ID NOS: 162  
; SEQ ID NO 26  
; LENGTH: 333  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-796-753-26

Query Match 54.9%; Score 39; DB 9; Length 333;  
Best Local Similarity 54.5%; Pred. No. 29;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWKNMRKV 11

Db 48 CYGWRNKGV 58

## RESULT 11

US-09-798-869-8  
; Sequence 8, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of caprine  
; OTHER INFORMATION: sequence)  
US-09-798-869-8

Query Match 53.5%; Score 38; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 2.4;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWKNMRKV 11

Db 3 CLRQWEMRKV 13

## RESULT 12

US-09-798-869-29  
; Sequence 29, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 29  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: BOVINE  
US-09-798-869-29

Query Match 53.5%; Score 38; DB 9; Length 15;  
Best Local Similarity 45.5%; Pred. No. 2.4;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWKNMRKV 11

|||:|:

Db 3 CFQWQRMKKL 13

## RESULT 13

JS-09-798-869-30  
; Sequence 30, Application US/09798869  
; Publication No. US2003002282A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 30  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: BOVINE  
JS-09-798-869-30

Query Match 53.5%; Score 38; DB 9; Length 15;  
Best Local Similarity 45.5%; Pred. No. 2.4;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQRMKKV 11

Db 3 CFQWQRMKKL 13

## RESULT 14

JS-09-864-761-47985  
; Sequence 47985, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aemica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 47985  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL096701.14  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6  
; OTHER INFORMATION: EST\_HUMAN HIT: AM294800.1, EVALUATE 1.00e-06  
US-09-864-761-47985

Query Match 53.5%; Score 38; DB 10; Length 21;  
Best Local Similarity 83.3%; Pred. No. 3.3;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWKR 6

Db 16 CFQWR 21

## RESULT 15

US-09-888-320-2  
; Sequence 2, Application US/09888320  
; Publication No. US20030013090A1  
; GENERAL INFORMATION:  
; APPLICANT: Barry III, Clifton E.  
; APPLICANT: DeBarber, Andrea E.  
; APPLICANT: Mdulii, Khisimuzi  
; APPLICANT: Bekker, Linda-Gail  
; APPLICANT: The Government of the United States of America  
; APPLICANT: as represented by The Secretary of the  
; APPLICANT: Department of Health and Human Services  
; TITLE OF INVENTION: Methods of Diagnosing Multidrug Resistant Tuberculosis  
; FILE REFERENCE: 015280-413100US  
; CURRENT APPLICATION NUMBER: US/09/888,320  
; CURRENT FILING DATE: 2001-06-22  
; PRIOR APPLICATION NUMBER: US 60/214,187  
; PRIOR FILING DATE: 2000-06-26  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 489  
; TYPE: PRT  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: wild-type Etas monooxygenase (Rv3854c, EthA)  
US-09-888-320-2

Query Match 53.5%; Score 38; DB 9; Length 489;  
Best Local Similarity 54.5%; Pred. No. 59;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQRMKKV 11

Db 253 CFQWQRMKKM 263



Search completed: February 21, 2003, 08:11:56  
Job time : 6.88372 secs

---

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

DM protein - protein search, using sw model

Run on: February 21, 2003, 07:48:01 ; Search time 10.6047 Seconds  
(without alignments)  
108.784 Million cell updates/sec

Title: US-09-743-107b-91

Perfect score: 71

Sequence: 1 CFQWXRNRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 73:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	67	94.4	711	1 TPHUL	Lactotransferrin p
2	50	70.4	708	2 J22323	Lactoferrin - goat
3	47	66.2	33	2 S2107	Lactoferrin - sheep
4	45	63.4	275	2 T22597	Hypothetical prote
5	41	57.7	206	2 H97451	Pyridoxamine 5',-ph
6	41	57.7	206	2 AB2670	Probable pyridoxam
7	41	57.7	208	2 AC3441	Hypothetical prote
8	41	57.7	511	2 AB0858	Lactoferrin precu
9	41	57.7	707	1 A28438	Hypothetical prote
10	40	56.3	584	2 C84325	Dynamin beta heavy
11	40	56.3	4568	2 T08030	Hypothetical prote
12	39	54.9	447	2 T18633	Hypothetical prote
13	39	54.9	759	2 G65066	Hypothetical prote
14	39	54.9	759	2 G72115	Hypothetical prote
15	38	53.5	60	2 A48396	Ribosomal protein
16	38	53.5	62	2 A41301	Ribosomal protein
17	38	53.5	267	2 A16703	Ribosomal protein
18	38	53.5	267	2 S77802	Hypothetical prote
19	38	53.5	298	2 AD2346	Hypothetical prote
20	38	53.5	393	2 B64639	3-Deoxy-manno-octu
21	38	53.5	489	2 C70655	Probable monooxyge
22	38	53.5	583	2 T01470	Diphosphate-fructo
23	37	52.1	205	2 E90094	26S proteasome SU
24	37	52.1	214	2 E07989	Vif protein - simi
25	37	52.1	274	2 B60950	Apolipoprotein B-1
26	37	52.1	357	2 T22879	Hypothetical prote
27	37	52.1	513	2 E86156	Tl4P4.7 protein -
28	37	52.1	518	2 B4514	Probable cytochrom
29	37	52.1	536	2 T24218	Hypothetical prote

## ALIGNMENTS

### RESULT 1

TPHUL

Lactotransferrin precursor [validated] - human

N/Alternate names: lactoferrin

C/Species: Homo sapiens (man)

C/Date: 31-Mar-1992 #sequence, revision 21-Nov-1997 #text change 08-Dec-2000

C/Accession: G01394; S11228; A45401; S10324; S15853; S20841; S07160; A61169; A31000; S74  
R/Cho, Y.

submitted to the EMBL Data Library, March 1994

A/Reference number: G06820

A/Accession: G01394

A/Status: preliminary; translated from GB/EMBL/DD8U

A/Molecule type: mRNA

A/Residues: 1-711 <CHO>  
A/Cross-references: EMBL:U07643; NID:G467236; PIDN:AAB60324.1; PID:G467237

R/Rey, M.W.; Woloshuk, S.L.; deBoer, H.A.; Pieper, F.R.

Nucleic Acids Res. 18, 5288, 1990

A/Title: Complete nucleotide sequence of human mammary gland lactoferrin.

A/Reference number: S11228; MUID:90384839; PMID:2402455

A/Accession: S11228

A/Molecule type: mRNA

A/Residues: 1-148,'T',150-422,'C',424-711 <REY>

A/Cross-references: EMBL:X53961; NID:G34415; PIDN:CAA37914.1; PID:G34416

R/Teng, C.T.; Liu, Y.; Yang, N.; Walmer, D.; Panella, T.

Mol. Endocrinol. 6, 1969-1981, 1992

A/Title: Differential molecular mechanism of the estrogen action that regulates lactoferrin

A/Reference number: A45401; MUID:93125571; PMID:1480183

A/Accession: A45401

A/Molecule type: DNA

A/Residues: 1-15 <TEN>

A/Cross-references: GB:S52659; NID:G263311; PIDN:AAB24877.1; PID:G263312

A/Experimental source: Placenta

A/Note: Sequence extracted from NCBI backbone (NCBIP.122202)

R/Powell, M.J.; Ogden, J.E.

Nucleic Acids Res. 18, 4013, 1990

A/Title: Nucleotide sequence of human lactoferrin cDNA.

A/Reference number: S10324; MUID:90326549; PMID:2374734

A/Accession: S10324

A/Molecule type: mRNA

A/Residues: 3-711 <POW>

A/Cross-references: EMBL:X52941; NID:G34411; PIDN:CAA37116.1; PID:G34412

R/Stowell, K.M.; Rado, T.A.; Funk, W.D.; Tweedie, J.W.

Biochem. J. 276, 349-355, 1991

A/Title: Expression of cloned human lactoferrin in baby hamster kidney cells.

A/Reference number: S15853; MUID:91264786; PMID:2049066

A/Accession: S15853

A/Status: nucleic acid sequence not shown; not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 20-31 <ST1>

A/Accession: S20841

A/Molecule type: protein

A/Residues: 20-28,'X',30-31 <ST2>

hypothetical prote  
GTP-binding regula  
hypothetical prote  
trichosyalin like  
hypothetical prote  
hypothetical prote  
F1511.22 [imported  
hypothetical prote  
60S ribosomal prot  
hypothetical prote  
hypothetical prote  
pre-pilin leader p  
33.3K hypothetical  
hypothetical prote  
hypothetical prote  
MHC class I histoc  
pectate lyase (EC

30 37 52.1 558 2 T17324  
31 37 52.1 846 2 S2418  
32 37 52.1 932 2 T28820  
33 37 52.1 1432 2 B85431  
34 36 50.7 99 2 AC1714  
35 36 50.7 99 2 AG1343  
36 36 50.7 124 2 C96582  
37 36 50.7 131 2 H91246  
38 36 50.7 159 2 T00407  
39 36 50.7 283 2 G8020  
40 36 50.7 287 2 E82546  
41 36 50.7 289 2 G86403  
42 36 50.7 349 2 T16882  
43 36 50.7 361 2 T29571  
44 36 50.7 365 2 I37477  
45 36 50.7 374 1 WZWCPC

R:Rado, T.A.; Wei, X.; Benz Jr., E.J.

Blood 70, 989-993, 1987  
A>Title: Isolation of lactoferrin cDNA from a human myeloid library and expression of mRNA  
A:Reference number: S07160; MUID:88001031; PMID:3477300

A:Accession: S07160

A:Molecule type: mRNA

A:Residues: 436-487, 'A', 489-711 <RAD>

A:Cross-references: EMBL:M18642; NID:AAA86665.1; PID:G386855

R:Panella, T.J.; Liu, Y.; Huang, A.T.; Teng, C.T.

Cancer Res. 51, 3037-3043, 1991

A>Title: Polymorphism and altered methylation of the lactoferrin gene in normal leukocytes

A:Reference number: A61169; MUID:91235214; PMID:1674448

A:Accession: A61169

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 3-701, 'SWKPNV' <PAN>

A:Experimental source: normal breast tissue

R:Metz-Boutigue, M.H.; Jolles, J.; Mazurier, J.; Schoentgen, F.; Legrand, D.; Spik, G.;

Eur. J. Biochem. 145, 659-666, 1984

A>Title: Human lactotransferrin: amino acid sequence and structural comparisons with other

A:Reference number: A31000; MUID:85076667; PMID:6510420

A:Accession: A31000

A:Molecule type: protein

A:Residues: 20-140, 142-169, 171-203, 'L', 205, 'K', 207-208, 'K', 210-385, 'Q', 387-391, 'W', 393-4

A>Note: this is the final paper in a series

R:Hoven, G.; Hoegdall, E.V.; Barkholt, V.; Norskov, L.

Eur. J. Biochem. 241, 303-308, 1996

A>Title: Lactoferrin: similarity to diamine oxidase and purification by aminoethyl affinity

A:Reference number: S74119; MUID:97054624; PMID:8898921

A:Accession: S74119

A:Molecule type: Protein

A:Residues: 'G', 23-24, 'R', 26-27, 'XX', 30-32 <HO>

A:Experimental source: neutrophil granulocytes

C:Genetics:

A:Gene: GDB:LTF

A:Cross-references: GDB:119368; OMIM:150210

A:Map position: 3q21-3q23

C:Superfamily: transferrin; transferrin repeat homology

C:Keywords: duplication; glycoprotein; iron binding; milk

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-711/Product: lactotransferrin #status experimental <MAT>

F:21-386/Domain: transferrin repeat homology <TRH>

F:360-699/Domain: transferrin repeat homology <TRH2>

F:29-65, 39-56, 135-218, 177-193, 190-201, 251-265, 503-697, 595-609/Disulfide bonds: #status e

F:157, 498/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:368-400, 378-391, 425-706, 447-669, 479-554, 513-527, 524-537, 647-652/Disulfide bonds: #stat

Query Match 94.4%; Score 67; DB 1; Length 711;

Best Local Similarity 91.7%; Pred. No. 0.00075;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWKNMKRV 12

Db 39 CFQWKNMKRV 50

RESULT 2

JC2323

lactoferrin - goat

C:Species: Capra aegagrus hircus (domestic goat)

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 07-May-1999

C:Accession: JC2323

R:Le Provost, F.; Nocart, M.; Guerin, G.; Martin, P.

Biochem. Biophys. Res. Commun. 203, 1324-1332, 1994

A>Title: Characterization of the goat lactoferrin cDNA: assignment of the relevant locus

A:Reference number: JC2323; MUID:94380047; PMID:8093048

A:Accession: JC2323

A:Molecule type: mRNA

A:Residues: 1-708 <LEP>

C:Superfamily: transferrin; transferrin repeat homology

C:Keywords: duplication; glycoprotein

F:359-886/Domain: transferrin repeat homology <TRH2>

F:252, 300, 387, 495, 564/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 70.4%; Score 50; DB 2; Length 708;

Best Local Similarity 53.6%; Pred. No. 0.72;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWKNMKRV 11

Db 38 CFQWKNMKRV 48

RESULT 3

S52107

lactoferrin - sheep (fragment)

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C:Date: 14-Jul-1995 #sequence\_revision 05-Dec-1997 #text\_change 05-Dec-1997

C:Accession: S52107

R:Qian, Z.Y.; Jolles, P.; Migliore-Samour, D.; Fiat, A.M.

Biochim. Biophys. Acta 1243, 25-32, 1995

A>Title: Isolation and characterization of sheep lactoferrin, an inhibitor of platelet

A:Reference number: S52107; MUID:95127729; PMID:7827104

A:Accession: S52107

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-33 <QIA>

C:Superfamily: transferrin; transferrin repeat homology

C:Keywords: duplication

Query Match 66.2%; Score 47; DB 2; Length 33;

Best Local Similarity 54.5%; Pred. No. 0.12;

Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWKNMKRV 11

Db 19 CFQWKNMKRV 29

RESULT 4

T22597

hypothetical protein F53H4.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000

C:Accession: T22597

R:Dobson, R.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19587

A:Accession: T22597

A>Status: preliminary; translated from GB/ENBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-275 <WIL>

A:Cross-references: EMBL:Z81089; PIDN:CAB03137.1; GSPDB:GN00028; CESP:F53H4.4

A:Experimental source: clone F53H4

C:Genetics:

A:Gene: CESP:F53H4.4

A:Map position: X

A:Introns: 67/1; 153/1

C:Superfamily: Caenorhabditis elegans hypothetical protein F53H4.4

Query Match 63.4%; Score 45; DB 2; Length 275;

Best Local Similarity 72.7%; Pred. No. 2.2;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 FQWKNMKRV 12

Db 262 FQWKNMKRV 272

RESULT 5

H97451

pyridoxamine 5'-phosphate oxidase (AF179611) [imported] - Agrobacterium tumefaciens (st

C:Species: Agrobacterium tumefaciens

C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002

C:Accession: H97451

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Goldman

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001  
 A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens  
 A;Reference number: A97359; PMID:11743194  
 A;Accession: H97451  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-206 <KUR>  
 A;Cross-references: GB:AB007869; PIDN:AAK86569.1; PID:G15155733; GSPDB:GN00169  
 C;Genetics:  
 A;Gene: AGR\_C1381  
 A;Map position: circular chromosome  
 C;Superfamily: pyridoxamine-phosphate oxidase

Query Match 57.7%; Score 41; DB 2; Length 206;  
 Best Local Similarity 58.3%; Pred. No. 8.2;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

DY 1 CFQWKNRMKVR 12  
 |||||  
 DB 88 CFHWKSLRQVR 99  
 |||||

RESULT 6  
 AB2670  
 A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58, Du  
 A;Reference number: AB2577; PMID:11743193  
 A;Accession: AB2670  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-206 <KUR>  
 A;Cross-references: GB:AB008689; PIDN:AAU41776.1; PID:G17739129; GSPDB:GN00186  
 A;Experimental source: strain C58 (Dupont)  
 C;Genetics:  
 A;Gene: pdxH  
 A;Map position: circular chromosome  
 C;Superfamily: pyridoxamine-phosphate oxidase

Query Match 57.7%; Score 41; DB 2; Length 206;  
 Best Local Similarity 58.3%; Pred. No. 8.2;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

DY 1 CFQWKNRMKVR 12  
 |||||  
 DB 88 CFHWKSLRQVR 99  
 |||||

RESULT 7  
 AG3441  
 A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
 A;Reference number: AG3252; PMID:11756688  
 A;Accession: AG3441  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-208 <KUR>  
 A;Cross-references: GB:AE008917; PIDN:AAU52698.1; PID:G17983525; GSPDB:GN00190

Query Match 57.7%; Score 41; DB 2; Length 206;  
 Best Local Similarity 58.3%; Pred. No. 8.2;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

DY 1 CFQWKNRMKVR 12  
 |||||  
 DB 88 CFHWKSLRQVR 99  
 |||||

RESULT 8  
 AB0858  
 A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhimurium  
 A;Reference number: AB0858  
 A;Accession: AB0858  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-511 <PAR>  
 A;Cross-references: GB:AL513382; PIDN:CAD06049.1; PID:G16504016; GSPDB:GN00176  
 C;Genetics:  
 A;Gene: STY3070

Query Match 57.7%; Score 41; DB 2; Length 511;  
 Best Local Similarity 58.3%; Pred. No. 20;  
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

DY 1 CFQWKNRMKVR 12  
 |||||  
 DB 350 CFAMDMMKAKVR 361  
 |||||

RESULT 9  
 A28438  
 A;Title: Lactoferrin precursor - mouse  
 A;Alternative names: lactotransferrin  
 C;Species: Mus musculus (house mouse)  
 C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C;Accession: A28438; A41205  
 R;Pentecost, B.T.; Teng, C.T.  
 J. Biol. Chem. 262, 10134-10139, 1987  
 A;Title: Lactotransferrin is the major estrogen inducible protein of mouse uterine secretory epithelial cells  
 A;Reference number: A92596; PMID:87280033; PMID:3611056  
 A;Accession: A28438  
 A;Molecule type: mRNA  
 A;Residues: 3-707 <PEN>  
 A;Cross-references: EMBL:J03298  
 R;Liu, Y.; Teng, C.T.  
 J. Biol. Chem. 266, 21880-21885, 1991  
 A;Title: Characterization of estrogen-responsive mouse lactoferrin promoter.  
 A;Reference number: A41205; PMID:92042099; PMID:1939212  
 A;Accession: A41205  
 A;Molecule type: DNA  
 A;Residues: 1-15 <LIU>  
 A;Cross-references: GB:M74778  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein  
 F;1-19/domain: signal sequence #status predicted <SIG>

A;Experimental source: strain 16M  
 C;Genetics:  
 A;Gene: BME11517  
 A;Map position: I  
 C;Superfamily: pyridoxamine-phosphate oxidase  
 C;Keywords: oxidoreductase

Query Match 57.7%; Score 41; DB 2; Length 208;  
 Best Local Similarity 58.3%; Pred. No. 8.3;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

DY 1 CFQWKNRMKVR 12  
 |||||  
 DB 90 CFHWKSLRQVR 101  
 |||||

RESULT 8  
 AB0858  
 A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhimurium  
 A;Reference number: AB0858  
 A;Accession: AB0858  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-511 <PAR>  
 A;Cross-references: GB:AL513382; PIDN:CAD06049.1; PID:G16504016; GSPDB:GN00176  
 C;Genetics:  
 A;Gene: STY3070

Query Match 57.7%; Score 41; DB 2; Length 511;  
 Best Local Similarity 58.3%; Pred. No. 20;  
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

DY 1 CFQWKNRMKVR 12  
 |||||  
 DB 350 CFAMDMMKAKVR 361  
 |||||

RESULT 9  
 A28438  
 A;Title: Lactoferrin precursor - mouse  
 A;Alternative names: lactotransferrin  
 C;Species: Mus musculus (house mouse)  
 C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C;Accession: A28438; A41205  
 R;Pentecost, B.T.; Teng, C.T.  
 J. Biol. Chem. 262, 10134-10139, 1987  
 A;Title: Lactotransferrin is the major estrogen inducible protein of mouse uterine secretory epithelial cells  
 A;Reference number: A92596; PMID:87280033; PMID:3611056  
 A;Accession: A28438  
 A;Molecule type: mRNA  
 A;Residues: 3-707 <PEN>  
 A;Cross-references: EMBL:J03298  
 R;Liu, Y.; Teng, C.T.  
 J. Biol. Chem. 266, 21880-21885, 1991  
 A;Title: Characterization of estrogen-responsive mouse lactoferrin promoter.  
 A;Reference number: A41205; PMID:92042099; PMID:1939212  
 A;Accession: A41205  
 A;Molecule type: DNA  
 A;Residues: 1-15 <LIU>  
 A;Cross-references: GB:M74778  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein  
 F;1-19/domain: signal sequence #status predicted <SIG>

F;20-707/Product: lactotransferrin #status predicted <MAT>  
 F;358-695/Domain: transferrin repeat homology <IRH2>  
 F;494/Binding site: carbonylurate (Asn) (covalent) #status predicted

Query Match 57.7%; Score 41; DB 1; Length 707;  
 Best Local Similarity 54.5%; Pred. No. 28;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWKNRKRKV 11  
 |||:|:|:|  
 Db 37 CLRQWNERKV 47

RESULT 10  
 C84325  
 hypothetical protein Vng1732c [imported] - Halobacterium sp. NRC-1  
 C;Species: Halobacterium sp. NRC-1  
 C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C;Accession: C84325  
 R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.;  
 ; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabil  
 Jung, K.H.; Alam, M.; Freitas, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
 A;Title: Genome sequence of Halobacterium species NRC-1.  
 A;Reference number: A84160; MUID:20504483; PMID:11016950  
 A;Accession: C84325  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-584 <STO>  
 A;Cross-references: GB:AE004437; NID:G10581192; PIDN:AAQ19967.1; GSPDB:GN00139  
 C;Genetics:  
 A;Gene: VNG1732C

Query Match 56.3%; Score 40; DB 2; Length 584;  
 Best Local Similarity 41.7%; Pred. No. 34;  
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWKNRKRKV 12  
 |||:|:|:|  
 Db 445 CFTWRKDMERKV 456

RESULT 11  
 T08030  
 dynein beta heavy chain - Chlamydomonas reinhardtii  
 C;Species: Chlamydomonas reinhardtii  
 C;Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 02-Feb-2001  
 C;Accession: T08030  
 R;Mitchell, D.R.; Brown, K.S.  
 J. Cell Sci. 107, 635-644, 1994  
 A;Title: Sequence analysis of the Chlamydomonas alpha and beta dynein heavy chain genes.  
 A;Reference number: Z16302; MUID:94274778; PMID:8006077  
 A;Accession: T08030  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-4568 <MIT>  
 A;Cross-references: EMBL:U02963; NID:G409965; PIDN:AAA19956.1; PID:G514215  
 A;Experimental source: strain 21gr  
 C;Genetics:  
 A;Gene: CDA4

A;Map position: IX  
 A;Introns: 48/2; 74/3; 105/3; 159/1; 225/3; 358/3; 514/3; 729/2; 919/3; 1004/3; 1115/3;  
 3334/3; 3686/3; 3892/3; 4240/3  
 C;Superfamily: dynein heavy chain, ciliary  
 C;Keywords: nucleotide binding; P-loop  
 F;1919-1926/Region: nucleotide-binding motif A (P-loop)  
 F;2202-2209/Region: nucleotide-binding motif A (P-loop)  
 F;2530-2537/Region: nucleotide-binding motif A (P-loop)

Query Match 56.3%; Score 40; DB 2; Length 4568;  
 Best Local Similarity 41.7%; Pred. No. 2.6e+02;  
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWKNRKRKV 12  
 |||:|:|:|  
 Db 1852 CFQWSQLRYIQ 1863

RESULT 12  
 T18633  
 hypothetical protein M18.8 - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
 C;Accession: T18633; T23799  
 R;Sims, M.  
 submitted to the EMBL Data Library, February 1996  
 A;Reference number: Z18999  
 A;Accession: T18633  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-447 <WIL>  
 A;Cross-references: EMBL:Z69634; PIDN:CAA93457.1; GSPDB:GN00022; CESP:M18.8  
 A;Experimental source: clone B0001  
 R;Steward, C.  
 submitted to the EMBL Data Library, January 1996  
 A;Reference number: Z19800  
 A;Accession: T23799  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-447 <W12>  
 A;Cross-references: EMBL:Z68507; PIDN:CAA92831.1; GSPDB:GN00022; CESP:M18.8  
 A;Experimental source: clone M18  
 C;Genetics:  
 A;Gene: CESP:M18.8  
 A;Map position: 4  
 A;Introns: 41/3; 137/1; 326/3; 434/2

Query Match 54.9%; Score 39; DB 2; Length 447;  
 Best Local Similarity 75.0%; Pred. No. 39;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 WKENMRKV 11  
 |||:|:|:|  
 Db 291 WKENLREV 298

RESULT 13  
 G86506  
 hypothetical protein CFJ0126 [imported] - Chlamydomonas pneumoniae (strain J138)  
 C;Species: Chlamydomonas pneumoniae, Chlamydia pneumoniae  
 C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Mar-2001  
 C;Accession: G86506  
 R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; I  
 Nucleic Acids Res. 28, 2311-2314, 2000  
 A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.  
 A;Reference number: A86491; MUID:20330349; PMID:10871362  
 A;Accession: G86506  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-759 <STO>  
 A;Cross-references: GB:BA000008; NID:G9978500; PIDN:BAA98337.1; GSPDB:GN00142  
 A;Experimental source: strain J138  
 C;Genetics:  
 A;Gene: CFJ0126

Query Match 54.9%; Score 39; DB 2; Length 759;  
 Best Local Similarity 66.7%; Pred. No. 66;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 QWKNRKRKV 11  
 |||:|:|:|  
 Db 488 QWKNLRDV 496

RESULT 14

72115  
 ypothetical protein CP0646 [imported] - Chlamydia pneumoniae (strains CWL029 and AR  
 ;Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
 ;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-May-2000  
 ;Accession: G72115, F81554  
 ;Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;  
 ature Genet. 21, 385-389, 1999  
 ;Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
 ;Reference number: A72000; NUID:99206606; PMID:10192388  
 ;Accession: G72115  
 ;Status: preliminary  
 ;Molecule type: DNA  
 ;Residues: 1-759 <ARN>  
 ;Cross-references: GB:AE001599; GB:AE001363; NID:94376387; PIDN:AAD18279.1; PID:9437639  
 ;Experimental source: strain CWL029  
 ;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,  
 C.; Dodson, R.; Gwinn, M.; Nelson, M.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
 Nucleic Acids Res. 28, 1397-1406, 2000  
 ;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
 ;Reference number: A81500; NUID:20150255; PMID:10684935  
 ;Accession: F81554  
 ;Status: preliminary  
 ;Molecule type: DNA  
 ;Residues: 1-759 <REA>  
 ;Cross-references: GB:AE002222; GB:AE002161; NID:97189553; PIDN:AAF38461.1; PID:9718956  
 ;Experimental source: strain AR39, HL cells  
 ;Genetics:  
 ;Gene: CP0126; CP0646

Query Match 54.9%; Score 39; DB 2; Length 759;  
 Best Local Similarity 66.7%; Pred. No. 66;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

2y 3 QWKNMKV 11  
 |||:|:|  
 2b 488 QWKNLRDV 496

RESULT 15  
 A48396  
 ribosomal protein L28 - Bacillus stearothermophilus  
 C:Species: Bacillus stearothermophilus  
 C:Date: 19-Nov-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Aug-1998  
 C:Accession: A48396  
 R;Kruft, V.; Kapp, U.; Wittmann-Liebold, B.  
 Biochimie 73, 855-860, 1991  
 A;Title: Characterization and primary structure of proteins L28, L33 and L34 from Bacill  
 A;Reference number: A48396; NUID:92075758; PMID:1742360  
 A;Accession: A48396  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-60 <KKU>  
 A;Note: sequence extracted from NCBI backbone (NCBIP:69662)  
 C;Superfamily: Escherichia coli ribosomal protein L28

Query Match 53.5%; Score 38; DB 2; Length 60;  
 Best Local Similarity 66.7%; Pred. No. 8.2;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 4 WKNMKV 12  
 |||:|:|  
 Db 27 WKANLQVR 35

Search completed: February 21, 2003, 08:02:46  
 Job time: 11.6047 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

DM protein - protein search, using sw model

Run on: February 21, 2003, 07:28:06 ; Search time 5.2093 Seconds  
(without alignments)  
95.544 Million cell updates/sec

Title: US-09-743-107B-91

Perfect score: 71

Sequence: 1 CFQKKNRKYR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	67	94.4	711	1 TRFL_HUMAN	P02788 homo sapien
2	50	70.4	708	1 TRFL_CAMDR	Q9tun0 camelus dro
3	50	70.4	708	1 TRFL_CAPHI	Q29477 capra hircu
4	42	59.2	695	1 TRFL_HORSE	Q77811 equus cabal
5	41	57.7	62	1 RL28_THETN	Q8r9u1 thermaanaer
6	41	57.7	707	1 TRFL_MOUSE	P08071 mus musculu
7	40	56.3	4568	1 DYHB_CHLRE	Q39565 chlamydomon
8	39	54.9	292	1 NLA_DROME	Q9xz18 drosophila
9	38	53.5	60	1 RL28_BACST	P23374 bacillus st
10	38	53.5	62	1 RL28_LISMO	Q92aj2 listeria mo
11	38	53.5	146	1 RPOB_LIBAF	P41187 liberibacte
12	38	53.5	267	1 Y125_MYCCA	P53661 mycoplasma
13	37	52.1	214	1 VIF_SIVS4	P12505 simian immu
14	37	52.1	783	1 YNR2_CAEEL	Q21988 caenorhabdi
15	36	50.7	365	1 IAX3_HUMAN	P30453 homo sapien
16	36	50.7	455	1 YXYL_CAEEL	Q19910 caenorhabdi
17	36	50.7	485	1 GLGA_BACST	O08328 bacillus st
18	36	50.7	502	1 C932_SOYBN	Q42798 glycine max
19	36	50.7	509	1 C931_SOYBN	Q42798 glycine max
20	36	50.7	510	1 C933_SOYBN	O81973 glycine max
21	36	50.7	528	1 CAX2_ARATH	Q38798 arabidopsis
22	36	50.7	530	1 CAX3_ARATH	P29402 arabidopsis
23	36	50.7	728	1 KDGI_ARATH	Q39017 arabidopsis
24	36	50.7	749	1 VP4_ROTGA	Q04916 rotavirus (
25	36	50.7	1135	1 PHYC_SORBI	P93528 sorghum bic
26	35	49.3	151	1 SYB2_RHIME	Q923q1 rhizobium m
27	35	49.3	215	1 VIF_HV25B	P12452 human immun
28	35	49.3	215	1 VIF_HV28T	P20878 human immun
29	35	49.3	343	1 SP2D_BACSU	F07372 bacillus su
30	35	49.3	344	1 SP2D_BACAM	P13251 bacillus am
31	35	49.3	351	1 NOV_CHICK	P28686 gallus gall
32	35	49.3	358	1 DDL_ENTHR	Q47827 enterococcu
33	35	49.3	398	1 YK13_CAEEL	P34337 caenorhabdi

## ALIGNMENTS

RESULT 1				
TRFL_HUMAN	STANDARD:	PRT:	711 AA	
ID	TRFL_HUMAN	Q16780; Q16785; Q16786; Q16789; Q00756; Q9H1Z3; Q96KZ4;		
AC	Q96KZ5;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Lactotransferrin precursor (Lactoferrin) [Contains: Lactoferrin A;			
DE	Lactoferrin B; Lactoferrin C].			
GN	LIF OR LF.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RX	MEDLINE=90384839; PubMed=2402455;			
RA	Rey M.W., Woloshuk S.L., de Boer H.A., Pieper F.R.;			
RT	"Complete nucleotide sequence of human mammary gland lactoferrin.";			
RL	Nucleic Acids Res. 18:5288-5288(1990).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Cho Y.Y.;			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Conneely O.M.;			
RL	Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Liang Q., Jimenez-Flores R., Richardson T.;			
RT	"Molecular cloning and sequence analysis of human lactoferrin.";			
RL	Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Bone marrow;			
RA	Wei X., Han J., Rado T.A.;			
RT	"Human neutrophil lactoferrin coding and 5' flanking region DNA			
RL	sequences.";			
RN	Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Cheng H., Chen X., Huan L.;			
RT	"cDNA cloning and sequence analysis of human lactoferrin.";			
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RN	[7]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Prostate;			
RA	Strausberg R.;			
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.			
RN	[8]			

Q9vfn2 drosophila  
Q9tj19 chlamydia m  
Q9ute7 schizosacch  
P35649 eikenella c  
P58306 thermoplasma  
Q9ulc6 homo sapien  
Q99up4 staphylococ  
P29328 ovis aries  
Q10353 schizosacch  
P43488 mus musculu  
P19506 simian immu  
Q55185 synechocyst

RP SEQUENCE OF 3-711 FROM N.A.  
RC TISSUE-Mammary Gland;  
RX MEDLINE=90326549; PubMed=2374734;  
RA Powell M.J., Ogden J.E.;  
RT "Nucleotide sequence of human lactoferrin cDNA.";  
RL Nucleic Acids Res. 18:4013-4013(1990).  
RN [9]  
RP SEQUENCE OF 20-711.  
RX MEDLINE=85076667; PubMed=6510420;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Schoentgen F.,  
RX Metz-Boutigue M.-H., Jolles J., Jolles P.;  
RA Legrand D., Spik G., Montreuil J., Jolles P.;  
RT "Human lactotransferrin: amino acid sequence and structural  
RT comparisons with other transferrins.";  
RL Eur. J. Biochem. 145:659-666(1984).  
RN [10]  
RP PRELIMINARY SEQUENCE OF 20-73; 134-171; 257-278 AND 347-530.  
RX MEDLINE=82046817; PubMed=6794640;  
RA Metz-Boutigue M.-H., Mazurier J., Jolles J., Spik G., Montreuil J.,  
RA Jolles P.;  
RT "The present state of the human lactotransferrin sequence. Study and  
RT alignment of the cyanogen bromide fragments and characterization of  
RT N- and C-terminal domains.";  
RL Biochim. Biophys. Acta 670:243-254(1981).  
RN [11]  
RP SEQUENCE OF 609-711.  
RX MEDLINE=82262043; PubMed=7049727;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,  
RA Jolles P.;  
RT "An 88 amino acid long C-terminal sequence of human  
RT lactotransferrin.";  
RL FEBS Lett. 142:107-110(1982).  
RN [12]  
RP SEQUENCE OF 436-711 FROM N.A.  
RX MEDLINE=88001031; PubMed=3477300;  
RA Rado T.A., Wei X., Benz E.J. Jr.;  
RT "Isolation of lactoferrin cDNA from a human myeloid library and  
RT expression of mRNA during normal and leukemic myelopoiesis.";  
RL Blood 70:989-993(1987).  
RN [13]  
RP SEQUENCE OF 237-711 FROM N.A.  
RA McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,  
RA Nham M., Parnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,  
RA Gnoj L., la Bastide M., Kaplan N., Greco T., Touchman J., Muzny D.,  
RA Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M., Porcel B.M.,  
RA Dragan Y., Giacalone J., Pae A., Powell E., Solinsky K.A., Desilva U.,  
RA Diaz-Perez S., Zhou X., Yu Y., Watanabe M., Doggett N., Garcia D.,  
RA Sagripanti J.L.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
RN [14]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND REVISIONS.  
RX MEDLINE=90064528; PubMed=2585506;  
RA Anderson B.F., Baker H.M., Norris G.E., Rice D.W., Baker E.N.;  
RT "Structure of human lactoferrin: crystallographic structure analysis  
RT and refinement at 2.8-A resolution.";  
RL J. Mol. Biol. 209:711-734(1989).  
RN [15]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RA Haridas M., Anderson B.F., Baker E.N.;  
RT "Structure of human diferric lactoferrin refined at 2.2-A  
RT resolution.";  
RL Acta Crystallogr. D 51:629-646(1995).  
RN [16]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 20-353.  
RX MEDLINE=97156796; PubMed=9003186;  
RA Nicholson H., Anderson B.F., Bland T., Shewry S.C., Tweedie J.W.,  
RA Baker E.N.;  
RT "Matogenesis of the histidine ligand in human lactoferrin: iron  
RT binding properties and crystal structure of the histidine-  
RT 253-->methionine mutant.";  
RL Biochemistry 36:341-346(1997).  
RN [17]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX MEDLINE=99190892; PubMed=10089347;  
RA Sun X.L., Baker H.M., Shewry S.C., Jameson G.B., Baker E.N.;  
RT "Structure of recombinant human lactoferrin expressed in Aspergillus  
RT awamori.";  
RL Acta Crystallogr. D 55:403-407(1999).  
RN [18]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=99192677; PubMed=10089508;  
RA Jameson G.B., Anderson B.F., Norris G.E., Thomas D.H., Baker E.N.;  
RT "Structure of human apolactoferrin at 2.0-A resolution. Refinement  
RT and analysis of ligand-induced conformational change.";  
RL Acta Crystallogr. D 54:1319-1335(1998).  
RN [19]  
RP CHARACTERIZATION OF LACTOFERROXINS.  
RX MEDLINE=911166929; PubMed=1369293;  
RA Tani F., Iio K., Chiba H., Yoshikawa M.;  
RT "Isolation and characterization of opioid antagonist peptides derived  
RT from human lactoferrin.";  
RL Agric. Biol. Chem. 54:1803-1810(1990).  
RN [20]  
RP VARIANTS THR-30 AND ARG-48.  
RX PubMed=9873069;  
RA Klintworth G.K., Sommer J.R., Obrian G., Han L., Ahmed M.N.,  
RA Qumsiyeh M.B., Lin P.-Y., Basti S., Reddy M.K., Kanai A., Hotta Y.,  
RA Sugar J., Kumaramanickavel G., Munier F., Schorderet D.F.,  
RA El Matiri L., Iwata F., Kalsner-Kupfer M., Nagata M., Nakayasu K.,  
RA Hejtmancik J.F., Teng C.T.;  
RT "Familial subepithelial corneal amyloidosis (gelatinous drop-like  
RT corneal dystrophy): exclusion of linkage to lactoferrin gene.";  
RL Mol. Vision 4:31-32(1998).  
CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
CC OF AN ANION, USUALLY BICARBONATE.  
CC -1- FUNCTION: LACTOFERROXINS A, B AND C HAVE OPIOID ANTAGONIST  
CC ACTIVITY. LACTOFERROXIN A SHOWS PREFERENCE FOR MU-RECEPTORS, WHILE  
CC LACTOFERROXIN B AND C HAVE SOMEWHAT HIGHER DEGREES OF PREFERENCE  
CC FOR KAPPA-RECEPTORS THAN FOR MU-RECEPTORS.  
CC -1- SUBUNIT: MONOMER.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X53961; CAA37914.1; -  
CC EMBL; U07643; AAB60324.1; -  
CC EMBL; X93150; AAA36159.1; -  
CC EMBL; M83202; AAA59811.1; -  
CC EMBL; M83205; AAA58656.1; -  
CC EMBL; M18642; AAA86665.1; -  
CC EMBL; AF332168; AAG48753.1; -  
CC EMBL; BC015822; AAH15822.1; -  
CC EMBL; BC015823; AAH15823.1; -  
CC EMBL; M73700; AAA59479.1; -  
CC EMBL; X52941; CAA37116.1; -  
CC EMBL; U95626; AAB57795.1; -  
CC PIR; S11228; TFHUL.  
CC PDB; 1LCF; 31-AUG-94.  
CC PDB; 1LCT; 31-OCT-93.  
CC PDB; 1LFG; 31-JUL-94.  
CC PDB; 1LFH; 31-OCT-93.  
CC PDB; 1LFI; 31-OCT-93.  
CC PDB; 1LGB; 31-AUG-94.  
CC PDB; 1LGC; 31-AUG-94.  
CC PDB; 1BKA; 08-NOV-96.  
CC PDB; 1DSN; 08-MAR-96.  
CC PDB; 1HSE; 12-MAR-97.  
CC PDB; 1VFD; 21-APR-97.



```
Query Match          94.4%; Score 67; DB 1; Length 711;
Best Local Similarity 91.7%; Pred. No. 0.0002;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWKENMKVR 12
    |||||
DB 39 CFQWQNRMKVR 50

RESULT 2
TRFL CAMDR          STANDARD;          PRT; 708 AA.
AC Q9TUM0; Q9MZS5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lactotransferrin precursor (Lactoferrin).
GN LTF.
OS Camelus dromedarius (Dromedary) (Arabian camel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.
OX NCBI_TaxID=9838;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Somali; TISSUE=Lactating mammary gland;
RA Kappler S.R., Ackermann M., Farah Z., Puhani Z.;
RT "Sequence analysis of camel (Camelus dromedarius) lactoferrin.";
RL Int. Dairy J. 9:481-486 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Paramasivam M., Srinivasan A., Singh R., Sahani M.S., Singh T.P.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE (BY SIMILARITY).
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ131674; CAB53387.1; -.
DR EMBL; AF165879; AAF82241.1; -.
DR HSSP; O77811; 1B1X.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin.2.
DR PRINTS; PR00422; TRANSFERRIN.
DR SMART; SM00094; TR_FER; 2.
DR PROSITE; PS00205; TRANSFERRIN_1; 2.
DR PROSITE; PS00206; TRANSFERRIN_2; 2.
DR PROSITE; PS00207; TRANSFERRIN_3; 2.
KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;
KW Signal.
FT SIGNAL. 1 19 BY SIMILARITY.
FT CHAIN 20 708 LACTOTRANSFERRIN.
FT REPEAT 20 363 1.
FT REPEAT 364 708 2.
FT DISULFID 28 64 BY SIMILARITY.
FT DISULFID 38 55 BY SIMILARITY.
FT DISULFID 134 217 BY SIMILARITY.
FT DISULFID 176 192 BY SIMILARITY.
FT DISULFID 189 200 BY SIMILARITY.
FT DISULFID 250 264 BY SIMILARITY.
FT DISULFID 367 399 BY SIMILARITY.
```

```
FT DISULFID 377 390 BY SIMILARITY.
FT DISULFID 424 703 BY SIMILARITY.
FT DISULFID 444 666 BY SIMILARITY.
FT DISULFID 476 551 BY SIMILARITY.
FT DISULFID 500 694 BY SIMILARITY.
FT DISULFID 510 524 BY SIMILARITY.
FT DISULFID 521 534 BY SIMILARITY.
FT DISULFID 592 606 BY SIMILARITY.
FT DISULFID 644 649 BY SIMILARITY.
FT METAL 79 79 IRON 1 (BY SIMILARITY).
FT METAL 111 111 IRON 1 (BY SIMILARITY).
FT METAL 211 211 IRON 1 (BY SIMILARITY).
FT METAL 272 272 IRON 1 (BY SIMILARITY).
FT METAL 414 414 IRON 2 (BY SIMILARITY).
FT METAL 452 452 IRON 2 (BY SIMILARITY).
FT METAL 545 545 IRON 2 (BY SIMILARITY).
FT METAL 614 614 IRON 2 (BY SIMILARITY).
FT BINDING 140 140 ANION (BY SIMILARITY).
FT BINDING 482 482 ANION (BY SIMILARITY).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 261 261 F -> S (IN REF. 2).
FT CONFLICT 304 304 G -> A (IN REF. 2).
FT CONFLICT 330 330 S -> P (IN REF. 2).
FT CONFLICT 492 494 LLS -> PLF (IN REF. 2).
FT CONFLICT 506 506 L -> F (IN REF. 2).
FT CONFLICT 609 609 A -> P (IN REF. 2).
FT CONFLICT 642 642 R -> Q (IN REF. 2).
SQ SEQUENCE 708 AA; 77211 MW; 0B0C175A0B69D430 CRC64;

Query Match          70.4%; Score 50; DB 1; Length 708;
Best Local Similarity 66.7%; Pred. No. 0.22;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQNRMKVR 12
    |||||
DB 38 CAQWQNRMKVR 49

RESULT 3
TRFL CAPHI          STANDARD;          PRT; 708 AA.
ID TRFL_CAPHI;
AC Q29477; Q29479;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lactotransferrin precursor (Lactoferrin).
GN LTF.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Lee T., Yu S., Kim S.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=94380047; PubMed=8093048;
RA le Provost F., Nocart M., Guerin G., Martin P.;
RT "Characterization of the goat lactoferrin cDNA. Assignment of the
RT relevant locus to bovine U12 syntenic group.";
RL Biochem. Biophys. Res. Commun. 203:1324-1332 (1994).
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
```



FT DISULFID 463 538  
 FT DISULFID 487 681  
 FT DISULFID 497 511  
 FT DISULFID 508 521  
 FT DISULFID 579 593  
 FT DISULFID 631 636  
 FT METAL 66  
 FT METAL 98  
 FT METAL 198  
 FT METAL 259  
 FT METAL 401  
 FT METAL 439  
 FT METAL 532  
 FT METAL 601  
 FT BINDING 127  
 FT BINDING 469  
 FT CARBOHYD 143  
 FT CARBOHYD 287  
 FT CARBOHYD 482  
 SQ SEQUENCE 695 AA; 75991 MW; 07BB84D50E1B165D CRC64;

Query Match 59.2%; Score 42; DB 1; Length 695;  
 Best Local Similarity 58.3%; Pred. No. 5.7;  
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

2y 1 CFQWKNNRKVR 12  
 : : : : :  
 25 CAKQFNKKVR 36

RESULT 5  
 RL28 THETN STANDARD; PRT; 62 AA.  
 AC Q8R9U1;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 50S ribosomal protein L28.  
 EN RPMB OR TTE1495.  
 OS Thermoaerobacter tengcongensis.  
 CC Bacteria; Firmicutes; Clostridia; Thermoaerobacteriales;  
 CC Thermoaerobacteriaceae; Thermoaerobacter.  
 CX NCBI\_taxid=119072;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=21992816; PubMed=11997336;  
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;  
 RT "A complete sequence of T. tengcongensis genome.";  
 RL Genome Res. 12:689-700(2002).  
 CC -!- SIMILARITY: BELONGS TO THE L28P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CX EMBL; AEO13107; AM24713.1; -  
 CX Ribosomal protein; Complete proteome.  
 CX SEQUENCE 62 AA; 7037 MW; CAADB605C81D495D CRC64;

Query Match 57.7%; Score 41; DB 1; Length 62;  
 Best Local Similarity 70.0%; Pred. No. 0.69;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

2y 3 QWKNNRKVR 12  
 : : : : :  
 27 RWKPNRKVR 36

RESULT 6  
 TRFL\_MOUSE  
 ID TRFL\_MOUSE STANDARD; PRT; 707 AA.  
 AC P08071; P70690; Q61799; Q922P2;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin).  
 GN LTF.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Uterus;  
 RC MEDLINE=87280033; PubMed=36110556;  
 RA Pentecost B.T., Teng C.T.;  
 RT "Lactotransferrin is the major estrogen inducible protein of mouse  
 RT uterine secretions.";  
 RL J. Biol. Chem. 262:10134-10139(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Uterus;  
 RA Morishima K.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 1-14 FROM N.A.  
 RX MEDLINE=92042099; PubMed=1939212;  
 RA Liu Y., Teng C.T.;  
 RT "Characterization of estrogen-responsive mouse lactoferrin promoter.";  
 RL J. Biol. Chem. 266:21880-21895(1991).  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY BICARBONATE.  
 CC -!- SUBUNIT: MONOMER.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
 CC -----  
 CX This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CX EMBL; J03298; AAA0525.1; -  
 CX EMBL; D88510; BAA13633.1; -  
 CX EMBL; BC006904; AAH06904.1; -  
 CX EMBL; M74778; AAA39427.1; -  
 CX PIR; A28438; A28438.  
 CX HSP; P02788; ICB6.  
 CX MGD; MG1:96837; Ltf.  
 CX InterPro; IPR001156; Transferrin.  
 CX Pfam; PF00405; transferrin; 2.  
 CX PRINTS; PR00422; TRANSFERRIN.  
 CX SMART; SMO0094; TR FER; 2.  
 CX PROSITE; PS00205; TRANSFERRIN\_1; 1.  
 CX PROSITE; PS00206; TRANSFERRIN\_2; 2.  
 CX PROSITE; PS00207; TRANSFERRIN\_3; 2.  
 CX Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 CX Signal.  
 CX SIGNAL 1 19 BY SIMILARITY.  
 CX FT CHAIN 20 707 LACTOTRANSFERRIN.  
 CX REPEAT 20 357 1.

```

FT REPEAT 358 707
FT BY SIMILARITY.
FT DISULFID 27 63
FT BY SIMILARITY.
FT DISULFID 37 54
FT BY SIMILARITY.
FT DISULFID 133 216
FT BY SIMILARITY.
FT DISULFID 178 191
FT BY SIMILARITY.
FT DISULFID 185 199
FT BY SIMILARITY.
FT DISULFID 249 263
FT BY SIMILARITY.
FT DISULFID 366 398
FT BY SIMILARITY.
FT DISULFID 376 389
FT BY SIMILARITY.
FT DISULFID 423 702
FT BY SIMILARITY.
FT DISULFID 443 665
FT BY SIMILARITY.
FT DISULFID 475 550
FT BY SIMILARITY.
FT DISULFID 499 693
FT BY SIMILARITY.
FT DISULFID 509 523
FT BY SIMILARITY.
FT DISULFID 520 533
FT BY SIMILARITY.
FT DISULFID 591 605
FT BY SIMILARITY.
FT DISULFID 643 648
FT BY SIMILARITY.
FT METAL 78 78
FT IRON 1 (BY SIMILARITY).
FT METAL 110 110
FT IRON 1 (BY SIMILARITY).
FT METAL 210 210
FT IRON 1 (BY SIMILARITY).
FT METAL 271 271
FT IRON 1 (BY SIMILARITY).
FT METAL 413 413
FT IRON 2 (BY SIMILARITY).
FT METAL 451 451
FT IRON 2 (BY SIMILARITY).
FT METAL 544 544
FT IRON 2 (BY SIMILARITY).
FT METAL 613 613
FT IRON 2 (BY SIMILARITY).
FT BINDING 139 139
FT ANION (POTENTIAL).
FT BINDING 481 481
FT ANION (POTENTIAL).
FT CARBOHYD 118 118
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 494 494
FT MR -> IOG (IN REF. 1).
FT CONFLICT 1 2
FT R -> Q (IN REF. 2).
FT CONFLICT 25 25
FT CONFLICT 82 82
FT M -> L (IN REF. 2).
FT CONFLICT 359 359
FT A -> T (IN REF. 2).
FT CONFLICT 382 382
FT A -> D (IN REF. 1).
FT CONFLICT 449 449
FT E -> G (IN REF. 2).
FT CONFLICT 629 629
FT L -> V (IN REF. 1).
SQ SEQUENCE 707 AA; 77865 MW, P26AE03404C19A8 CRC64;

Query Match 57.7%; Score 41; DB 1; Length 707;
Best Local Similarity 54.5%; Pred. No. 8.7;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 CFQWKNMKRV 11
Db 37 CLRQWENMKV 47

RESULT 7
DYHB CHLRE
ID DYHB CHLRE STANDARD; PRT; 4568 AA.
AC Q3955;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Dynein beta chain, flagellar outer arm.
GN ODA4 OR ODA-4 OR SUP1.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
CX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=21GI;
RX MEDLINE=94274778; PubMed=8006077;
RA Mitchell D.R., Brown K.S.;
RT "Sequence analysis of the Chlamydomonas alpha and beta dynein heavy
chain genes.";
RL J. Cell Sci. 107:635-644 (1994).
CC -J- FUNCTION: FORCE GENERATING PROTEIN OF EUKARYOTIC CILIA AND
CC FLAGELLA. PRODUCES FORCE TOWARDS THE MINUS ENDS OF MICROTUBULES.
CC DYNEN HAS ATPASE ACTIVITY.
CC -I- SUBUNIT: CONSISTS OF AT LEAST 3 HEAVY CHAINS (ALPHA, BETA AND
GAMMA), 2 INTERMEDIATE CHAINS AND 8 LIGHT CHAINS.

```

```

CC -I- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U02963; AAA19956.1;
CC InterPro; IPR004273; Dynein heavy.
CC Pfam; PF03028; Dynein heavy; 1.
CC Motor protein; Microtubules; Dynein; ATP-binding; Flagella;
KW Coiled coil.
FT DOMAIN 277 293 COILED COIL (POTENTIAL).
FT DOMAIN 1158 1175 COILED COIL (POTENTIAL).
FT DOMAIN 1372 1400 COILED COIL (POTENTIAL).
FT DOMAIN 1614 1650 COILED COIL (POTENTIAL).
FT DOMAIN 1778 1825 COILED COIL (POTENTIAL).
FT DOMAIN 2017 2045 MICROTUBULE-BINDING (POTENTIAL).
FT DOMAIN 2831 2848 COILED COIL (POTENTIAL).
FT DOMAIN 3106 3162 COILED COIL (POTENTIAL).
FT DOMAIN 3339 3425 COILED COIL (POTENTIAL).
FT DOMAIN 3648 3728 COILED COIL (POTENTIAL).
FT NP_BIND 1919 1926 ATP (POTENTIAL).
FT NP_BIND 2202 2209 ATP (POTENTIAL).
FT NP_BIND 2530 2537 ATP (POTENTIAL).
FT NP_BIND 2879 2886 ATP (POTENTIAL).
SQ SEQUENCE 4568 AA; 519961 MW, 9A9A5393C7C36AE7 CRC64;

Query Match 56.3%; Score 40; DB 1; Length 4568;
Best Local Similarity 41.7%; Pred. No. 92;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 CFQWKNMKRV 12
Db 1852 CFQWQQLRYIQ 1863

RESULT 8
NLA DROME
ID NLA DROME STANDARD; PRT; 292 AA.
AC Q9XZL8; Q9V391;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nebula protein.
GN NLA OR CG6072.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA McCormick A.V., Goldberg M.L.;
RT "Gene required for elongation of meiosis I spindle in Drosophila
females.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams W.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brannon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos F.V., Berman B.P., Bhandari D., Bolshakov S.,

```



```

QY 4 WKRMKVR 12
DB 29 WKANLQVR 37

RESULT 11
RPOB LIBAF STANDARD; PRT; 146 AA.
AC P41187;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA-directed RNA polymerase beta chain (EC 2.7.7.6) (Transcriptase
beta chain) (RNA polymerase beta subunit) (Fragment).
GN RPOB.
OS Liberibacter africanus (Liberibacter africanus).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Candidatus Liberibacter.
OX NCBI_TaxID=34020;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Nelspruit;
RA Planet P., Jagoueix S., Bove J.M., Garnier M.;
RT "Detection and characterization of the African Citrus Greening
Liberibacter by amplification, cloning and sequencing of the rplKJL-
rpoBC operon.";
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
[RNA] (N).
CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
BETA' CHAIN.
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC
CC EMBL; U08675; AAA19557.1; -
CC InterPro; IPR001572; RNA_pol_B.
CC Pfam; PF00562; RNA_pol_B; 1.
CC PROSITE; PS01165; RNA_POL BETA; PARTIAL.
KW Transferase; Transcription; DNA-directed RNA polymerase.
FT NON TER 146
SQ SEQUENCE 146 AA; 16731 MW; 815D6EDD68FD9943 CRC64;

Query Match 53.5%; Score 38; DB 1; Length 146;
Best Local Similarity 50.0%; Pred. No. 5.7;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWKNRNX 10
DB 10 CVQWSRGARK 19

RESULT 12
Y125 MYCCA STANDARD; PRT; 267 AA.
AC P53651;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE Hypothetical protein (Fragment).
OS Mycoplasma capricolum.
OC Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;
OC Entomoplasmataceae.

QY 4 WKRMKVR 12
DB 29 WKANLQVR 37

NCBI_TaxID=2095;
(1)
SEQUENCE FROM N.A.
STRAIN=AFCC 27343 / KID;
MEDLINE=96059641; PubMed=7476192;
Bork P., Ouzounis C., Casari G., Schneider R., Sander C.,
Dolan M., Gilbert W., Gillevet P.M.;
"Exploring the Mycoplasma capricolum genome: a minimal cell reveals
its physiology.";
Mol. Microbiol. 16:955-967(1995).
-1- SIMILARITY: BELONGS TO THE COF/YEHA/YIDA/YIGL (E. COLI) / YCSE/YXEH
(B. SUBTILIS) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC
CC EMBL; Z33006; CAA83589.1; -
CC InterPro; IPR001454; Hlgnaase/hydrilase.
CC InterPro; IPR000150; Hypothet_cof.
CC Pfam; PF00702; Hydrolase; 1.
CC PROSITE; PS01228; COF 1; 1.
CC PROSITE; PS01229; COF 2; 1.
KW Hypothetical protein.
FT NON TER 267
SQ SEQUENCE 267 AA; 30425 MW; D5912DD5B39A8451 CRC64;

Query Match 53.5%; Score 38; DB 1; Length 267;
Best Local Similarity 58.3%; Pred. No. 11;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWKNRNX 12
DB 158 CFGKKNRQMR 169

RESULT 13
VIF_SIVS4 STANDARD; PRT; 214 AA.
AC P12505;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DE Virion infectivity factor (SOR protein) (Q protein).
GN VIF.
OS Simian immunodeficiency virus (F236/smH4 isolate) (sooty mangabey).
OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11737;
RN (1)
RP SEQUENCE FROM N.A.
MEDLINE=89262053; PubMed=2786147;
Hirsch V.M., Olmstead R.A., Murphy-Corb M., Purcell R.H.,
Johnson P.R.;
"An African primate lentivirus (SIVsm) closely related to HIV-2.";
Nature 339:389-392(1989).
-1- FUNCTION: DETERMINES VIRUS INFECTIVITY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X14307; CAA32484.1; -
CC PIR; S07989; S07989.
CC HIV; X14307; VIFSSMMH4.
CC InterPro; IPR000475; Viral_infect.

```

```
DR Pfam; PF00559; Vif; 1.
DR PRINTS; PR00349; VIRIONINFECT.
DR ProDom; PD000063; Viral_infect; 1.
DR AIDS.
SQ SEQUENCE 214 AA; 25140 MW; 9BCE884EC454BF3D CRC64;
Query Match 52.1%; Score 37; DB 1; Length 214;
Best Local Similarity 75.0%; Pred.No. 13;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
2Y 3 QWKRNRK 10
2b 170 QWRNRK 177
RESULT 14
YNR2_CABEL STANDARD; PRT; 783 AA.
AC Q21988;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein R13G10.2 in chromosome III.
EN R13G10.2.
DS Caenorhabditis elegans.
DC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
DC Rhabditidae; Peloderinae; Caenorhabditis.
DX NCBI_TaxID=6239;
EN [1]
SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gardner A.E.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Durbin R.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DC -!- COFACTOR: FAD (POTENTIAL).
DC -!- SIMILARITY: BELONGS TO THE FLAVIN MONOAMINE OXIDASE FAMILY.
DC This SWISS-PROT entry is copyright. It is produced through a collaboration
DC between the Swiss Institute of Bioinformatics and the EMBL outstation -
DC the European Bioinformatics Institute. There are no restrictions on its
DC use by non-profit institutions as long as its content is in no way
DC modified and this statement is not removed. Usage by and for commercial
DC entities requires a license agreement (See http://www.isb-sib.ch/announce/
DC or send an email to license@isb-sib.ch).
DC -----
DC EMBL; Z35602; CA84671.2; -
DC WormPep; R13G10.2; CE25088.
DC InterPro; IPR002937; Amino_oxidase.
DC Pfam; PF01593; Amino_oxidase; 1.
DC Hypothetical protein; Oxidoreductase; Flavoprotein; FAD.
DC NP_BIND 311 366 FAD (ADP PART) (POTENTIAL).
PT SEQUENCE 783 AA; 88799 MW; 8D087E9E464DC908 CRC64;
Query Match 52.1%; Score 37; DB 1; Length 783;
Best Local Similarity 50.0%; Pred.No. 50;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
2Y 1 CFQWKRNRKVR 12
2b 540 CIDWGRDNRKVR 551
RESULT 15
LA34_HUMAN STANDARD; PRT; 365 AA.
LC P30453; P30454;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE HLA class I histocompatibility antigen, AW-34(A-10) alpha chain
```

```
DE precursor.
GN HLA-A OR HLA-A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
EN [1]
SEQUENCE FROM N.A. (A*3401/A*3402).
MEDLINE=93056508; PubMed=1491115;
RA Madrigal J.A., Belich M.P., Hildebrand W.H., Benjamin R.J.,
RA Little A.-M., Zemmour J., Ennis P.D., Ward F.E., Petzl-Erler M.L.,
RA Martell R.W., du Toit E.D., Parham P.;
PT "Distinctive HLA-A,B antigens of black populations formed by
PT interallelic conversion.";
RL J. Immunol. 149:3411-3415(1992).
EN [2]
SEQUENCE FROM N.A. (A*3401/A*3402).
MEDLINE=93235211; PubMed=8475492;
RA Madrigal J.A., Hildebrand W.H., Belich M.P., Benjamin R.J.,
RA Little A.-M., Zemmour J., Ennis P.D., Ward F.E., Petzl-Erler M.L.,
RA du Toit E.D., Parham P.;
PT "Structural diversity in the HLA-A10 family of alleles: correlations
PT with serology.";
RL Tissue Antigens 41:72-80(1993).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -!- POLYMORPHISM: THE FOLLOWING ALLELES OF AW-34 ARE KNOWN: A*3401
CC (AW-34.1) AND A*3402 (AW34.2). THE SEQUENCE SHOWN IS THAT OF
CC A*3401.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X61704; CA43873.1; -
CC EMBL; X61705; CA43874.1; -
CC PIR; S16767; S16767.
CC PIR; S16771; S16771.
CC HSP; O19673; IHSB.
CC MIM; 142800; -
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR003597; Ig_cl.
CC InterPro; IPR001039; MHC_I.
CC Pfam; PF00047; Ig; 1.
CC ProDom; PD000050; MHC_I; 1.
CC SMART; SM00407; IGcl; 1.
CC PROSITE; PS00230; IG_MHC; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal; Polymorphism.
FT SIGNAL 1 24
FT CRAIN 25 365
FT -----
FT HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT AW-34(A-10) ALPHA CHAIN.
FT EXTRACELLULAR ALPHA-1.
FT EXTRACELLULAR ALPHA-2.
FT EXTRACELLULAR ALPHA-3.
FT CONNECTING PEPTIDE.
FT -----
FT CYTOPLASMIC TAIL.
FT N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT BY SIMILARITY.
FT BY SIMILARITY.
FT I -> V (IN A*3402).
FT /FTID=VAR_004379.
FT K -> N (IN A*3402).
FT /FTID=VAR_004380.
FT R -> I (IN A*3402).
FT /FTID=VAR_004381.
FT -----
FT VARIANT 121 121
FT VARIANT 121 121
```

FT VARIANT 129 129 P -> S (IN A\*3402).  
 /FTId=VAR 004382.  
 FT VARIANT 138 138 Q -> R (IN A\*3402).  
 /FTId=VAR 004383.  
 FT VARIANT 180 180 W -> L (IN A\*3402).  
 /FTId=VAR 004384.  
 FT VARIANT 312 312 L -> I (IN A\*3402).  
 /FTId=VAR 004385.  
 SQ SEQUENCE 365 AA; 41055 MW; 063BF63E686E01F6 CRC64;

Query Match 50.7%; Score 36; DB 1; Length 365;  
 Best Local Similarity 66.7%; Pred. NC. 34;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 WKRNRKVR 12  
 Db 84 WDRNRKVK 92

Search completed: February 21, 2003, 07:51:38  
 Job time : 6.2093 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:44:43 ; Search time 21.6744 Seconds  
(without alignments)  
114.078 Million cell updates/sec

Title: US-09-743-107B-91

Perfect score: 71

Sequence: 1 CFQWKNRKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 21:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_mhc:\*

9: sp\_mhc:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteriaph:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	62	87.3	711	4 Q8TCD2	Q8tcd2 homo sapien
2	58	81.7	38	4 Q9UCY5	Q9ucy5 homo sapien
3	47	66.2	33	6 Q9TR80	Q9tr80 ovis aries
4	45	63.4	275	5 Q93780	Q93780 caenorhabdi
5	41	57.7	62	16 Q8R9U1	Q8r9ul thermocaaer
6	41	57.7	205	16 Q986A0	Q986a0 rhizobium l
7	41	57.7	206	16 Q8UHC2	Q8uhc2 agrobacteri
8	41	57.7	208	16 Q92RH8	Q92rh8 rhizobium m
9	41	57.7	208	16 Q8YFK3	Q8yfk3 brucella me
10	41	57.7	511	16 Q82462	Q82462 salmonella
11	40	56.3	81	15 Q90863	Q90863 human immun
12	40	56.3	148	10 Q9XHP1	Q9xhpi sesamum ind
13	40	56.3	274	4 Q96M21	Q96m21 homo sapien
14	40	56.3	584	17 Q9HPA3	Q9hpa3 halobacteri
15	39	54.9	273	2 Q31090	Q31090 rhizobium l
16	39	54.9	279	16 Q8XSE2	Q8xse2 ralstonia s

17	39	54.9	447	5	017549	017549 caenorhabdi
18	39	54.9	550	11	Q9JUZ5	Q9jjz5 mus musculus
19	39	54.9	759	16	Q9Z955	Q9z955 chlamydia p
20	39	54.9	864	5	O62582	O62582 encephalito
21	39	54.9	864	5	Q8SRG3	Q8srg3 encephalito
22	39	54.9	864	5	Q8SQI6	Q8sqi6 encephalito
23	38	53.5	81	15	Q90884	Q90884 human immun
24	38	53.5	91	15	Q77855	Q77855 human immun
25	38	53.5	105	10	Q9XFD5	Q9xids oryza sativ
26	38	53.5	207	10	Q9SML1	Q9sml1 cicer ariet
27	38	53.5	240	10	Q9SML2	Q9sml2 cicer ariet
28	38	53.5	253	12	Q69541	Q69541 horseradish
29	38	53.5	298	16	Q8YP77	Q8yp77 anabaena sp
30	38	53.5	306	4	Q8TAX2	Q8tax2 homo sapien
31	38	53.5	329	2	Q9F7Y4	Q9f7y4 salmonella
32	38	53.5	341	11	Q8R2A4	Q8r2a4 mus musculus
33	38	53.5	372	10	O81653	O81653 hemerocalli
34	38	53.5	393	10	Q9ZTP0	Q9ztp0 oryza sativ
35	38	53.5	393	16	O25611	O25611 helicobacte
36	38	53.5	402	10	Q9ZRH8	Q9zrh8 oryza sativ
37	38	53.5	466	4	Q9NUS2	Q9nus2 homo sapien
38	38	53.5	489	16	P96223	P96223 mycobacteri
39	38	53.5	499	10	Q9FXI1	Q9fxi1 cicer ariet
40	38	53.5	560	10	Q9FTT0	Q9ftt0 oryza sativ
41	38	53.5	570	10	Q89487	Q89487 zea mays (m
42	38	53.5	583	10	O81437	O81437 arabidopsis
43	38	53.5	662	12	Q9QU30	Q9qu30 ttv-like mi
44	38	53.5	666	5	Q9WIDS	Q9wids drosophila
45	38	53.5	866	10	Q9FHI9	Q9fhi9 arabidopsis

## ALIGNMENTS

### RESULT 1

Q8TCD2 PRELIMINARY; PRT; 711 AA.  
AC Q8TCD2; 01-JUN-2002 (T-EMBLrel. 21, Created)  
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)  
DE Lactotransferrin.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PROSTATE;  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC022347; AAH22347.1; -;  
SQ SEQUENCE 711 AA; 78327 MW; 1B9C7E097C45FAF CRC64;

Query Match 87.3%; Score 62; DB 4; Length 711;  
Best Local Similarity 90.9%; Pred. No. 0.0074;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWKNRKVR 11  
Db 39 CFQWKNRKVR 49

### RESULT 2

Q9UCY5 PRELIMINARY; PRT; 38 AA.  
ID Q9UCY5  
AC Q9UCY5; 01-MAY-2000 (T-EMBLrel. 13, Created)  
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)  
DE Lactoferrin homolog (Fragment)  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96081613; PubMed=8551695;  
RA Sato I.;  
RT "Characterization of the 84-kDa protein with ABH activity in human  
RT seminal plasma.";  
RL Jpn. J. Legal Med. 49:281-293(1995).  
DR HSSP; P02788; 1BKRA.  
DR InterPro; IPR001156; Transferrin.  
DR Pfam; PF00405; transferrin; 1.  
SQ SEQUENCE 38 AA; 4459 MW; 0402F490B5BDDDEB CRC64;

Query Match 81.7%; Score 58; DB 4; Length 38;  
Best Local Similarity 90.9%; Pred. No. 0.0019;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 FQWKNMRKVR 12  
| | | | | | | | | |  
Db 21 FQWKNMRKVR 31

RESULT 3  
Q9TR80 PRELIMINARY; PRT; 33 AA.  
AC Q9TR80;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE Lactoferrin (Fragment).  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=95127729; PubMed=7827104;  
RA Qian Z.Y., Jolles P., Migliore-Samur D., Fiat A.M.;  
RL Biochim. Biophys. Acta 1243:25-32(1995).  
DR HSSP; O77698; 1CE2.  
DR InterPro; IPR001156; Transferrin.  
DR Pfam; PF00405; transferrin; 1.  
SQ SEQUENCE 33 AA; 3914 MW; D1904CAE15A73961 CRC64;

Query Match 66.2%; Score 47; DB 6; Length 33;  
Best Local Similarity 54.5%; Pred. No. 0.16;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWKNMRKV 11  
| | | | | | | | | |  
Db 19 CFQWKNMRKL 29

RESULT 4  
Q93780 PRELIMINARY; PRT; 275 AA.  
AC Q93780;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE F53H4.4 protein.  
GN F53H4.4  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dobson R.;  
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]

RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C.elegans: A platform for  
RT investigating biology.";  
RL Science 282:2012-2018(1998).  
DR EMBL; Z81089; CAB03137.1; -;  
SQ SEQUENCE 275 AA; 31716 MW; 9D69524FFC704DB2 CRC64;

Query Match 63.4%; Score 45; DB 5; Length 275;  
Best Local Similarity 72.7%; Pred. No. 3.1;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQWKNMRKVR 12  
| | | | | | | | | |  
Db 262 FQWKNMRKTR 272

RESULT 5  
Q8R9U1 PRELIMINARY; PRT; 62 AA.  
AC Q8R9U1;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Ribosomal protein L28.  
GN RPMB OR TTE1495.  
OS Thermoanaerobacter tengcongensis  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
OC Thermoanaerobacteriales; Thermoanaerobacteriaceae; Thermoanaerobacter.  
OX NCBI\_TaxID=119072;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MB47 / JCM11007;  
RX MEDLINE=21992816; PubMed=11997336;  
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
RA Tan H., Chen R., Wang J., Yu J., Yang H.;  
RT "A complete sequence of T. tengcongensis genome.";  
RL Genome Res. 12:689-700(2002).  
DR EMBL; AE013107; AAM24713.1; -;  
KW Complete proteome.  
SQ SEQUENCE 62 AA; 7037 MW; CAADB605C81D495D CRC64;

Query Match 57.7%; Score 41; DB 16; Length 62;  
Best Local Similarity 70.0%; Pred. No. 3.6;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 QWKNMRKVR 12  
| | | | | | | | | |  
Db 27 RWKPNIRKVR 36

RESULT 6  
Q986A0 PRELIMINARY; PRT; 205 AA.  
AC Q986A0;  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Pyridoxamine 5'-phosphate oxidase.  
GN MLL7454.  
OS Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Phyllobacteriaceae; Mesorhizobium.  
OX NCBI\_TaxID=381;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MAFF303099;  
RX MEDLINE=21082930; PubMed=11214968;  
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,  
Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,

RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,  
 RA Takeuchi C., Yamada M., Tabata S.,  
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
 RT Mesorhizobium loti.",  
 RL DNA Res. 7:331-338(2000).  
 DR EMBL; AP03011; BAB53553.1; -.  
 DR InterPro; IPR000659; Pyridox oxidase.  
 DR Pfam; PF01243; Pyridox oxidase; 1.  
 DR ProDom; PD006312; Pyridox oxidase; 1.  
 DR TIGRfam; TIGR00558; pdxH; 1.  
 DR PROSITE; PS01064; PYRIDOX\_OXIDASE; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 205 AA; 23300 MW; 0BADE4CD312327EA CRC64;

Query Match 57.7%; Score 41; DB 16; Length 205;

Best Local Similarity 58.3%; Pred. No. 12; Mismatches 4; Indels 0; Gaps 0;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

# RESULT 7

ID Q8UHC2 PRELIMINARY; PRT; 206 AA.  
 AC Q8UHC2;  
 DT 01-JUN-2002 (TRENBLrel. 21, Created)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)  
 DE Pyridoxamine 5'-phosphate oxidase.  
 EN PDHX OR ATU0760 OR AGR C 1381.  
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.  
 OX NCBI\_TaxID=176299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21608550; PubMed=11743193;  
 RA Wood D.W., Setubal J.C., Kaul R., Marks D.E., Kitajima J.P.,  
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,  
 RA Chen Y., Paulsen I.F., Eisen J.A., Karp P.D., Bovee D. Sr.,  
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,  
 RA Kutyavina T., Levy R., Li M.-J., McClelland E., Palmieri A.,  
 RA Raymond C., Rouse G., Saenphumchak C., Wu Z., Romero P., Gordon D.,  
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
 RA Nester E.W.;  
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens  
 RT C58.",  
 RL Science 294:2317-2323(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21608551; PubMed=11743194;  
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
 RA Grollio B., Gorman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,  
 RA Houmel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,  
 RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,  
 RA Flanagan C., Crowell C., Gursion J., Lono C., Sear C., Strub G.,  
 RA Cielo C., Slater S.;  
 RT "Genome sequence of the plant pathogen and biotechnology agent  
 RT Agrobacterium tumefaciens C58.",  
 RL Science 294:2323-2328(2001).  
 DR EMBL; AE009043; AAL41776.1; -.  
 DR EMBL; AE008009; AAK86569.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 206 AA; 23720 MW; 3BE488AE5307C0C1 CRC64;

# Query Match

Best Local Similarity 57.7%; Score 41; DB 16; Length 206;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWKNNRKVR 12  
 |||||:|  
 DB 88 CFHWKSLRRQVR 99

# RESULT 8

ID Q92RH8 PRELIMINARY; PRT; 206 AA.  
 AC Q92RH8;  
 DT 01-DEC-2001 (TRENBLrel. 19, Created)  
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
 DE Probable pyridoxamine 5'-phosphate oxidase (PNP/EMP oxidase) protein  
 DE (EC 1.4.3.5).  
 GN PDHX OR R00895 OR SMC00069.  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Sinorhizobium.  
 OX NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=1021;  
 RX MEDLINE=21396507; PubMed=11481430;  
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,  
 RA Pohl T., Portetle D., Puehler A., Fumelle B., Ramsperger U.,  
 RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;  
 RT "Analysis of the chromosome sequence of the legume symbiont  
 RT Sinorhizobium meliloti strain 1021.",  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 DR EMBL; AL591785; CAC45467.1; -.  
 DR InterPro; IPR000659; Pyridox oxidase.  
 DR Pfam; PF01243; Pyridox oxidase; 1.  
 DR ProDom; PD006312; Pyridox oxidase; 1.  
 DR TIGRfam; TIGR00558; pdxH; 1.  
 DR PROSITE; PS01064; PYRIDOX\_OXIDASE; UNKNOWN 1.  
 KW Oxidoreductase; Complete proteome.  
 SQ SEQUENCE 206 AA; 23900 MW; A2DB74229DACA97A CRC64;

Query Match 57.7%; Score 41; DB 16; Length 206;

Best Local Similarity 58.3%; Pred. No. 12; Mismatches 4; Indels 0; Gaps 0;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

# RESULT 9

ID Q8YFK3 PRELIMINARY; PRT; 208 AA.  
 AC Q8YFK3;  
 DT 01-MAR-2002 (TRENBLrel. 20, Created)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
 DE Probable pyridoxamine 5'-phosphate oxidase (EC 1.4.3.5).  
 GN BME1157.  
 OS Brucella melitensis.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Brucellaceae; Brucella.  
 OX NCBI\_TaxID=29459;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=16M / ATCC 23456 / BIOTYPE 1;  
 RX MEDLINE=20020109; PubMed=11756688;  
 RA DelVecchio V.G., Kapral V., Redkar R.J., Patra G., Mujar C., Los T.,  
 RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,  
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,  
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,  
 RA Haselkorn R., Kyripides N., Overbeek R.;  
 RT "The genome sequence of the facultative intracellular pathogen  
 RT Brucella melitensis.",

```
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448 (2002).
DR EMBL; AE009587; AAL52598.1; -.
DR InterPro; IPR000659; Pyridox_Oxidase.
DR Pfam; PF01243; Pyridox_oxidase; 1.
DR ProDom; PD006312; Pyridox_oxidase; 1.
DR TIGRFAMs; TIGR00558; pdxH; 1.
DR PROSITE; PS01064; PYRIDOX_OXIDASE; 1.
DR OXIDOREDUCTASE; Complete proteome.
SQ SEQUENCE 208 AA; 23866 MW; CB1F50BC9612DE28 CRC64;

Query Match 57.7%; Score 41; DB 16; Length 208;
Best Local Similarity 58.3%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWKNNMKVR 12
   ||| ||| |||
Db 90 CFHWKSLRRQVR 101

RESULT 10
Q8Z462 08Z462 PRELIMINARY; PRT; 511 AA.
AC 08Z462;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein STY3070.
GN STY3070.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davies P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar typhi CT18."
RL Nature 413:848-852 (2001).
DR EMBL; AL627276; CAD06049.1; -.
SQ SEQUENCE 511 AA; 58126 MW; E2DDDD124E10D178B CRC64;

Query Match 57.7%; Score 41; DB 16; Length 511;
Best Local Similarity 58.3%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CFQWKNNMKVR 12
   ||| ||| |||
Db 350 CFADWNNKAKVR 361

RESULT 11
Q90863 090863 PRELIMINARY; PRT; 81 AA.
AC 090863;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Glycoprotein gp120 (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=NUS182;
RA Brandful J.A.M., Ampofo W.K., Janssens W., Adu-Sarkodie Y.,
RA Apegyei F., Anyomi F., Aidoo S., Barnor J.S., Yamamoto N.,
RA Ishikawa K., Sata T., Kurata T.;
RT "Genetic and phylogenetic analysis of HIV-1 strains from Southern
RT Ghana."
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ225659; CAA12541.1; -.
DR InterPro; IPR000777; GPI20.
DR Pfam; PF00516; GPI20; 1.
DR AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 81
SQ SEQUENCE 81 AA; 9138 MW; 2D43DCD554295572 CRC64;

Query Match 56.3%; Score 40; DB 15; Length 81;
Best Local Similarity 66.7%; Pred. No. 7.1;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 QWKNNMKV 11
   :||| :|||
Db 69 EWKENLKV 77

RESULT 12
Q9XHP1 09XHP1 PRELIMINARY; PRT; 148 AA.
AC 09XHP1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE 2S albumin.
OS Sesamum indicum (Oriental sesame) (Gingelly).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Lamiales; Pedaliaceae; Sesamum.
OX NCBI_TaxID=4182;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TAINAN 1;
RX MEDLINE=20074970; PubMed=10606554;
RA Tai S.S.K., Wu L.S.H., Chen E.C.F., Tzen J.T.C.;
RT "Molecular cloning of 11S globulin and 2S albumin, the two major seed
RT storage proteins in sesame."
RL J. Agric. Food Chem. 47:4932-4938 (1999).
DR EMBL; AF091841; AAD42943.1; -.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR000617; Napin.
DR InterPro; IPR001768; Try/amyI_inhbr.
DR Pfam; PF00234; tryp_alpha_amiI; 1.
DR PRINTS; PR00496; NAFIN.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 148 AA; 17524 MW; BA46B033BA13E3DE CRC64;

Query Match 56.3%; Score 40; DB 10; Length 148;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWKNNMR 9
   ||| ||| |||
Db 54 CMQWMSMR 62

RESULT 13
Q96M21 096M21 PRELIMINARY; PRT; 274 AA.
ID Q96M21;
AC Q96M21;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CDNA FJ332891 fig. clone TEST12004929.
OS Homo sapiens (Human).
```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=TESTIS;  
 RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,  
 RA Arita M., Mueashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,  
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,  
 RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,  
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,  
 RA Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,  
 RA Sugano S., Nagahori K., Maehuo Y., Nagai K., Isogai T.;  
 RT "NEDO human cDNA sequencing project";  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AK057433; BAB71433.1; -;  
 SQ SEQUENCE 274 AA; 30083 MW; 1DD43654D4135B2F CRC64;

Query Match 56.3%; Score 40; DB 4; Length 274;  
 Best Local Similarity 50.0%; Pred. No. 25;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

2Y 1 CFQWKNMKVR 12  
 |||:|:|:  
 2b 66 CFQWGWGVYLR 77

RESULT 14  
 29HPA3 PRELIMINARY; PRT; 584 AA.  
 ID Q9HPA3  
 AC Q9HPA3;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Vng1732c.  
 DE Vng1732c.  
 EN Vng1732c.  
 OS Halobacterium sp. (strain NRC-1).  
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
 CC Halobacteriaceae; Halobacterium.  
 DX NCBI\_TaxID=64091;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20504483; PubMed=11016950;  
 RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,  
 RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsen V., Sbrogna J.,  
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo V.A.,  
 RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angewine C.M., Dale H.,  
 RA Isenbarger T.A., Peck R.F., Fohlstroder M., Spudich J.L., Jung K.-H.,  
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
 RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;  
 RT "Genome sequence of Halobacterium species NRC-1";  
 UL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
 DR EMBL; AE005078; AAG19967.1; -;  
 DR InterPro; IPR001646; 5peptide\_repeat.  
 DR InterPro; IPR001622; K-channel\_pore.  
 DR Pfam; PF00805; Pentapeptide; 2.  
 KW Complete proteome.  
 SQ SEQUENCE 584 AA; 65151 MW; 21BF5D5F0486CCC6 CRC64;

Query Match 56.3%; Score 40; DB 17; Length 584;  
 Best Local Similarity 41.7%; Pred. No. 54;  
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

2Y 1 CFQWKNMKVR 12  
 |||:|:|:  
 2b 445 CFQWKNMKVR 456

RESULT 15  
 231090 PRELIMINARY; PRT; 273 AA.  
 ID O31090  
 AC O31090;

DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
 DE Hypothetical 31.0 kDa protein.  
 OS Rhizobium leguminosarum (biovar viciae).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.  
 OX NCBI\_TaxID=387;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VF39;  
 RA Ivashina T.V., Sadykov M.R., Kanapin A.A., Ksenzenko V.N.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VF39;  
 RX MEDLINE=99113394; PubMed=9914965;  
 RA Sadykov M.R., Ivashina T.V., Kanapin A.A., Shlyapnikov M.G.,  
 RA Ksenzenko V.N.;  
 RT "Structural and functional organization of the exopolysaccharide  
 RT biosynthesis genes in Rhizobium leguminosarum bv. viciae VF39";  
 RL Mol. Biol. (Mosk) 32:797-804(1998).  
 DR EMBL; AF028810; AAB88891.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 273 AA; 30986 MW; F195D2FEF7AD44D3 CRC64;

Query Match 54.9%; Score 39; DB 2; Length 273;  
 Best Local Similarity 60.0%; Pred. No. 37;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 QWKNMKVR 12  
 :|||:|:  
 Db 245 RWLNLRKLR 254

Search completed: February 21, 2003, 08:00:44  
 Job time : 21.6744 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:37:21 ; Search time 28.093 Seconds  
(without alignments)  
56.918 Million cell updates/sec

Title: US-09-743-107B-92

Perfect score: 71

Sequence: 1 CFQWKENMKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002.\*  
1: /SID32/gcgdata/geneseq/geneseq-embl/AA1980.DAT.\*  
2: /SID32/gcgdata/geneseq/geneseq-embl/AA1981.DAT.\*  
3: /SID32/gcgdata/geneseq/geneseq-embl/AA1982.DAT.\*  
4: /SID32/gcgdata/geneseq/geneseq-embl/AA1983.DAT.\*  
5: /SID32/gcgdata/geneseq/geneseq-embl/AA1984.DAT.\*  
6: /SID32/gcgdata/geneseq/geneseq-embl/AA1985.DAT.\*  
7: /SID32/gcgdata/geneseq/geneseq-embl/AA1986.DAT.\*  
8: /SID32/gcgdata/geneseq/geneseq-embl/AA1987.DAT.\*  
9: /SID32/gcgdata/geneseq/geneseq-embl/AA1988.DAT.\*  
10: /SID32/gcgdata/geneseq/geneseq-embl/AA1989.DAT.\*  
11: /SID32/gcgdata/geneseq/geneseq-embl/AA1990.DAT.\*  
12: /SID32/gcgdata/geneseq/geneseq-embl/AA1991.DAT.\*  
13: /SID32/gcgdata/geneseq/geneseq-embl/AA1992.DAT.\*  
14: /SID32/gcgdata/geneseq/geneseq-embl/AA1993.DAT.\*  
15: /SID32/gcgdata/geneseq/geneseq-embl/AA1994.DAT.\*  
16: /SID32/gcgdata/geneseq/geneseq-embl/AA1995.DAT.\*  
17: /SID32/gcgdata/geneseq/geneseq-embl/AA1996.DAT.\*  
18: /SID32/gcgdata/geneseq/geneseq-embl/AA1997.DAT.\*  
19: /SID32/gcgdata/geneseq/geneseq-embl/AA1998.DAT.\*  
20: /SID32/gcgdata/geneseq/geneseq-embl/AA1999.DAT.\*  
21: /SID32/gcgdata/geneseq/geneseq-embl/AA2000.DAT.\*  
22: /SID32/gcgdata/geneseq/geneseq-embl/AA2001.DAT.\*  
23: /SID32/gcgdata/geneseq/geneseq-embl/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	71	100.0	12	21 AAY78091	Human lactoferrin
2	71	100.0	12	21 AAY78092	Human lactoferrin
3	67	94.4	12	21 AAY78038	Human lactoferrin
4	67	94.4	12	21 AAY78045	Human lactoferrin
5	67	94.4	12	21 AAY78047	Human lactoferrin
6	67	94.4	12	21 AAY78084	Human lactoferrin
7	67	94.4	13	21 AAY78037	Human lactoferrin
8	67	94.4	13	21 AAY78048	Human lactoferrin
9	67	94.4	13	21 AAY78049	Human lactoferrin
10	67	94.4	14	21 AAY78036	Human lactoferrin

11	67	94.4	14	21	AAY78050	Human lactoferrin
12	67	94.4	14	21	AAY78051	Human lactoferrin
13	67	94.4	15	17	AA98554	Peptide for anti-u
14	67	94.4	15	21	AAY78035	Human lactoferrin
15	67	94.4	15	21	AAY78062	Human lactoferrin
16	67	94.4	15	21	AAY78063	Human lactoferrin
17	67	94.4	16	21	AAY78031	Human lactoferrin
18	67	94.4	16	21	AAY78064	Human lactoferrin
19	67	94.4	16	21	AAY78065	Human lactoferrin
20	67	94.4	17	21	AAY78034	Human lactoferrin
21	67	94.4	17	21	AAY78066	Human lactoferrin
22	67	94.4	17	21	AAY78067	Human lactoferrin
23	67	94.4	18	15	AA65352	Human lactoferrin
24	67	94.4	18	17	AAW13397	Advanced glycosyla
25	67	94.4	18	21	AAY78033	Human lactoferrin
26	67	94.4	19	21	AAY68867	Amino acid sequenc
27	67	94.4	19	21	AAY78032	Human lactoferrin
28	67	94.4	20	13	AA21810	Anti microbial pep
29	67	94.4	20	14	AA44841	Lactoferrin-relate
30	67	94.4	20	15	AA48530	Lactoferrin derive
31	67	94.4	20	15	AA48531	Lactoferrin derive
32	67	94.4	20	15	AA57461	Lactoferrin derive
33	67	94.4	20	15	AA57462	Lactoferrin derive
34	67	94.4	20	16	AA84698	Bovine lactoferrin
35	67	94.4	20	16	AA84699	Bovine lactoferrin
36	67	94.4	20	16	AA80263	Anti-parasitic lac
37	67	94.4	20	16	AA80264	Anti-parasitic lac
38	67	94.4	20	17	AA98553	Peptide for anti-u
39	67	94.4	20	17	AA91852	Lactoferrin-derive
40	67	94.4	20	17	AAW03045	Lactoferrin-derive
41	67	94.4	20	17	AA930607	Lactoferrin-derive
42	67	94.4	20	17	AA937621	Lactoferrin-derive
43	67	94.4	20	17	AA937622	Lactoferrin-derive
44	67	94.4	20	18	AAW26150	Lactoferrin deriva
45	67	94.4	20	18	AAW14036	Anti-parasitic pep

#### ALIGNMENTS

RESULT 1  
AAY78091  
ID AAY78091 standard; Peptide; 12 AA.  
XX  
AC AAY78091;  
XX  
DT 25-APR-2000 (first entry)  
XX  
DE Human lactoferrin derived peptide SEQ ID NO:91.  
XX  
KW Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200001730-A1.  
XX  
PD 13-JAN-2000.  
XX  
PF 06-JUL-1999; 99WO-SE01230.  
XX  
PR 06-JUL-1998; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
XX  
PA (ASCI-) A+ SCI INVEST AB.  
XX  
PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX  
DR WPI; 2000-147388/13.



CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.

XX SQ Sequence 12 AA;

Query Match 94.4%; Score 67; DB 21; Length 12;

Best Local Similarity 91.7%; Pred. No. 9.7e-05; Mismatches 0; Indels 0; Gaps 0;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWRNRKVR 12

||||:|||||

Db 1 CFQWRNRKVR 12

#### RESULT 4

AAV78046 ID AAY78046 standard; Peptide; 12 AA.

XX AC AAY78046;

XX DT 25-APR-2000 (first entry)

DE Human lactoferrin derived peptide SEQ ID NO:46.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.

XX OS Homo sapiens.

OS Synthetic.

XX WC200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattheby-Baltzer I, Baltzer L, Dolphin GT;  
XX WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections, -  
PT inflammations and tumors and for use in infant formula food -

XX PS Claim 15; Page 35; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.

CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.

XX SQ Sequence 12 AA;

Query Match 94.4%; Score 67; DB 21; Length 12;

Best Local Similarity 91.7%; Pred. No. 9.7e-05; Mismatches 0; Indels 0; Gaps 0;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWRNRKVR 12

||||:|||||

Db 1 CFQWRNRKVR 12

#### RESULT 5

AAV78047 ID AAY78047 standard; Peptide; 12 AA.

XX AC AAY78047;

XX DT 25-APR-2000 (first entry)

DE Human lactoferrin derived peptide SEQ ID NO:47.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.

XX OS Homo sapiens.

OS Synthetic.

XX PN WC200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattheby-Baltzer I, Baltzer L, Dolphin GT;  
XX WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections, -  
PT inflammations and tumors and for use in infant formula food -

XX PS Claim 18; Page 73; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.

CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.

XX SQ Sequence 12 AA;

Query Match 94.4%; Score 67; DB 21; Length 12;

Best Local Similarity 91.7%; Pred. No. 9.7e-05; Mismatches 0; Indels 0; Gaps 0;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;



QY 1 CFQWKNRKR 12  
 DB 1 CFQWQNRKR 12

## RESULT 6

AA78084  
 ID AAY78084 standard; Peptide; 12 AA.

AC AAY78084;  
 XX  
 DT 25-APR-2000 (first entry)  
 XX Human lactoferrin derived peptide SEQ ID NO:84.  
 DE Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 OS  
 PN WO200001730-A1.  
 XX  
 PD 13-JAN-2000.  
 XX  
 PF 06-JUL-1999; 99WO-SE01230.  
 XX  
 PR 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 PA (ASCII-) A+ SCI INVEST AB.  
 XX  
 PI Hanson LA, Mattesby-Baltzer I, Baltzer L, Dolphin GT;  
 XX  
 DR WPI; 2000-147388/13.  
 XX  
 PT New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX  
 PS Claim 22; Page 36; 102pp; English.  
 XX  
 CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 SQ Sequence 12 AA;

Query Match 94.4%; Score 67; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 9.7e-05;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWKNRKR 12  
 DB 1 CFQWQNRKR 12

## RESULT 7

AA78037

ID AAY78037 standard; Peptide; 13 AA.  
 XX  
 AC AAY78037;  
 XX  
 DT 25-APR-2000 (first entry)  
 XX  
 DE Human lactoferrin derived peptide SEQ ID NO:37.  
 XX  
 KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 OS  
 PN WO200001730-A1.  
 XX  
 PD 13-JAN-2000.  
 XX  
 PF 06-JUL-1999; 99WO-SE01230.  
 XX  
 PR 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 PA (ASCII-) A+ SCI INVEST AB.  
 XX  
 PI Hanson LA, Mattesby-Baltzer I, Baltzer L, Dolphin GT;  
 XX  
 DR WPI; 2000-147388/13.  
 XX  
 PT New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX  
 PS Claim 12; Page 70; 102pp; English.  
 XX  
 CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 SQ Sequence 13 AA;

Query Match 94.4%; Score 67; DB 21; Length 13;  
 Best Local Similarity 91.7%; Pred. No. 0.00011;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWKNRKR 12  
 DB 2 CFQWQNRKR 13

## RESULT 8

AA78048  
 ID AAY78048 standard; Peptide; 13 AA.

XX  
 AC AAY78048;  
 XX  
 DT 25-APR-2000 (first entry)  
 XX  
 DE Human lactoferrin derived peptide SEQ ID NO:48.

```

XX Human, lactoferrin; modification; infection; inflammation; tumour;
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;
KW urinary tract infection; colitis; Candida infection; fungicidal;
KW bactericidal; preservative.
XX Homo sapiens.
OS Synthetic.
OS
XX WO200001730-A1.
XX 13-JAN-2000.
XX 06-JUL-1999; 99WO-S01230.
XX 06-JUL-1998; 98SE-0002441.
XX 17-JUL-1998; 98SE-0002562.
XX 29-DEC-1998; 98SE-0004614.
XX (ASCI-) A+ SCI INVEST AB.
XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;
XX WPI; 2000-147388/13.
XX New peptides used for treatment and prevention of infections,
XX inflammations and tumors and for use in infant formula food
XX Claim 15; Page 74; 102pp; English.
XX
XX AAY78001 to AAY78100 represent peptides having sequences based on human
XX lactoferrin. The peptides are taken up in the intestine through
XX binding to specific lactoferrin receptors and are then transported
XX through the circulation. A medicinal product of the peptide or fragment
XX can be used for treating and/or prevention of infections (such as
XX urinary tract infections, colitis, and Candida infection on a mucosal
XX membrane), inflammations and/or tumours. The peptides can also be used
XX in food stuffs such as infant formula food. The peptides are also
XX fungicidal and bactericidal and may also be used as preservatives.
XX Even though native human lactoferrin have been shown to have desired
XX anti-inflammatory anti-infectious and anti-tumoural properties they
XX cannot be used clinically on a broad basis because of high production
XX costs. Therefore, provision of peptides based on lactoferrin would
XX enable them to be used for the same purposes as lactoferrin at lower
XX cost.
XX
XX Sequence 13 AA;
XX
XX Query Match 94.4%; Score 67; DB 21; Length 13;
XX Best Local Similarity 91.7%; Pred. No. 0.00011;
XX Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 CFQWRNMRKVR 12
XX ||||:|||||
XX Db 2 CFQWRNMRKVR 13
XX
XX RESULT 9
XX AAY78049
XX ID AAY78049 standard; Peptide; 13 AA.
XX
XX AC AAY78049;
XX
XX DT 25-APR-2000 (first entry)
XX
XX Human lactoferrin derived peptide SEQ ID NO:49.
XX
XX Human; lactoferrin; modification; infection; inflammation; tumour;
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;
KW urinary tract infection; colitis; Candida infection; fungicidal;
KW bactericidal; preservative.
XX Homo sapiens.
XX

```

```

OS Synthetic.
XX WO200001730-A1.
XX 13-JAN-2000.
XX 06-JUL-1999; 99WO-S01230.
XX 06-JUL-1998; 98SE-0002441.
XX 17-JUL-1998; 98SE-0002562.
XX 29-DEC-1998; 98SE-0004614.
XX (ASCI-) A+ SCI INVEST AB.
XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;
XX WPI; 2000-147388/13.
XX New peptides used for treatment and prevention of infections,
XX inflammations and tumors and for use in infant formula food
XX Claim 18; Page 74; 102pp; English.
XX
XX AAY78001 to AAY78100 represent peptides having sequences based on human
XX lactoferrin. The peptides are taken up in the intestine through
XX binding to specific lactoferrin receptors and are then transported
XX through the circulation. A medicinal product of the peptide or fragment
XX can be used for treating and/or prevention of infections (such as
XX urinary tract infections, colitis, and Candida infection on a mucosal
XX membrane), inflammations and/or tumours. The peptides can also be used
XX in food stuffs such as infant formula food. The peptides are also
XX fungicidal and bactericidal and may also be used as preservatives.
XX Even though native human lactoferrin have been shown to have desired
XX anti-inflammatory anti-infectious and anti-tumoural properties they
XX cannot be used clinically on a broad basis because of high production
XX costs. Therefore, provision of peptides based on lactoferrin would
XX enable them to be used for the same purposes as lactoferrin at lower
XX cost.
XX
XX Sequence 13 AA;
XX
XX Query Match 94.4%; Score 67; DB 21; Length 13;
XX Best Local Similarity 91.7%; Pred. No. 0.00011;
XX Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 CFQWRNMRKVR 12
XX ||||:|||||
XX Db 2 CFQWRNMRKVR 13
XX
XX RESULT 10
XX AAY78036
XX ID AAY78036 standard; Peptide; 14 AA.
XX
XX AC AAY78036;
XX
XX DT 25-APR-2000 (first entry)
XX
XX Human lactoferrin derived peptide SEQ ID NO:36.
XX
XX Human; lactoferrin; modification; infection; inflammation; tumour;
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;
KW urinary tract infection; colitis; Candida infection; fungicidal;
KW bactericidal; preservative.
XX Homo sapiens.
XX Synthetic.
XX WO200001730-A1.
XX 13-JAN-2000.
XX 06-JUL-1999; 99WO-S01230.
XX

```

XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 PA (ASCI-) A+ SCI INVEST AB.  
 XX

PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX

XX WPI; 2000-147388/13.  
 XX

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX

PS Claim 12; Page 69; 102pp; English.  
 XX

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX

SQ Sequence 14 AA;  
 XX

Query Match 94.4%; Score 67; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 0.00011;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWKNMKVR 12  
 DB 3 CFQWKNMKVR 14

RESULT 11

AAV78050  
 ID AAY78050 standard; Peptide; 14 AA.

XX

AC AAY78050;

XX 25-APR-2000 (first entry)

DT Human lactoferrin derived peptide SEQ ID NO:50.

XX

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX

OS Homo sapiens.

OS Synthetic.

XX WO200001730-A1.

PN 13-JAN-2000.

XX

PF 06-JUL-1999; 98WO-SE01230.

XX

PR 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

PA

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX

PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX  
 DR WPI; 2000-147388/13.  
 XX  
 PT New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX  
 PS Claim 15; Page 75; 102pp; English.  
 XX

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX

SQ Sequence 14 AA;  
 XX

Query Match 94.4%; Score 67; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 0.00011;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWKNMKVR 12  
 DB 3 CFQWKNMKVR 14

RESULT 12

AAV78051  
 ID AAY78051 standard; Peptide; 14 AA.

XX

AC AAY78051;

XX 25-APR-2000 (first entry)

DT Human lactoferrin derived peptide SEQ ID NO:51.

XX

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX

OS Homo sapiens.

OS Synthetic.

XX WO200001730-A1.

PN 13-JAN-2000.

XX

PF 06-JUL-1999; 98WO-SE01230.

XX

PR 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

PA

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX

XX WPI; 2000-147388/13.  
 XX

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX

PS Claim 18; Page 75; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 14 AA;

Query Match 94.4%; Score 67; DB 21; Length 14;  
 Best Local Similarity 91.7%; Pred. No. 0.00011;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRNMRKVR 12  
 ||||:|||||  
 Db 3 CFQWQRNMRKVR 14

RESULT 13  
 AAR98554  
 ID AAR98554 standard; Peptide; 15 AA.  
 XX AC AAR98554;  
 XX DT 12-NOV-1996 (first entry)  
 XX DE Peptide for anti-ulcer agent.  
 XX KW anti-ulcer agent; low toxicity; stable; heat-resistant.  
 XX OS Synthetic.  
 XX PN JP08143468-A.  
 XX PD 04-JUN-1996.  
 XX PF 17-NOV-1994; 94JP-0283869.  
 XX PR 17-NOV-1994; 94JP-0283869.  
 XX PA (MORG ) MORINAGA MILK IND CO LTD.  
 XX DR WPI; 1996-318857/32.  
 XX PT Anti-ulcer agent contg. peptide - has low toxicity, is  
 PT heat-resistant and water-soluble  
 XX PS Claim 1; Page 11; 11pp; Japanese.  
 XX CC AAR98531-54 are peptides used in an anti-ulcer agent. The agent is low  
 CC in toxicity, is heat-resistant and stable in aqueous soln.. It can be  
 CC administered orally and be produced in large amounts.

XX SQ Sequence 15 AA;

Query Match 94.4%; Score 67; DB 17; Length 15;  
 Best Local Similarity 91.7%; Pred. No. 0.00012;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRNMRKVR 12  
 ||||:|||||  
 Db 2 CFQWQRNMRKVR 13

RESULT 14  
 AAY78035  
 ID AAY78035 standard; Peptide; 15 AA.  
 XX AC AAY78035;  
 XX DT 25-APR-2000 (first entry)  
 XX DE Human lactoferrin derived peptide SEQ ID NO:35.  
 XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX OS Homo sapiens.  
 OS Synthetic.  
 XX PN WO200001730-A1.  
 XX PD 13-JAN-2000.  
 XX PF 06-JUL-1999; 99WO-SE01230.  
 XX PR 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX PA (ASCI-) A+ SCI INVEST AB.  
 XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX DR WPI; 2000-147398/13.  
 XX PT New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX PS Claim 12; Page 69; 102pp; English.

XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 15 AA;

Query Match 94.4%; Score 67; DB 21; Length 15;  
 Best Local Similarity 91.7%; Pred. No. 0.00012;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQRNMRKVR 12  
 ||||:|||||  
 Db 4 CFQWQRNMRKVR 15

RESULT 15  
 AAY78062  
 ID AAY78062 standard; Peptide; 15 AA.  
 XX AC AAY78062;

XX 25-APR-2000 (first entry)  
DT Human lactoferrin derived peptide SEQ ID NO:62.  
XX  
DE Human; lactoferrin; modification; infection; inflammation; tumour;  
XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
XX WO200001730-A1.  
FN  
XX  
XX 13-JAN-2000.  
PD  
XX  
XX 06-JUL-1999; 99WO-SB01230.  
FF  
XX  
XX 06-JUL-1998; 98SE-0002441.  
PR  
XX 17-JUL-1998; 98SE-0002562.  
PR  
XX 29-DEC-1998; 98SE-0004614.  
PR  
XX (ASCI-) A+ SCI INVEST AB.  
DA  
XX  
XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
FI  
XX WPI; 2000-147388/13.  
XX  
XX New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
PT  
XX  
XX Claim 15; Page 81; 102pp; English.  
PS  
XX  
XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.  
XX  
SQ Sequence 15 AA;

Query Match 94.4%; Score 67; DB 21; Length 15;  
Best Local Similarity 91.7%; Pred. No. 0.00012;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWKNMKVR 12  
Db 4 CFQWQNNMKVR 15

Search completed: February 21, 2003, 07:56:44  
Job time : 28.093 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:50:40 ; Search time 8.93023 Seconds  
(without alignments)  
39.537 Million cell updates/sec

Title: US-09-743-107B-92

Perfect score: 71

Sequence: 1 CFQWKNRNVKVR 12

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	67	94.4	18	1	US-08-204-487-3
2	67	94.4	18	2	US-08-485-948-8
3	67	94.4	18	2	US-08-628-380-8
4	67	94.4	18	2	US-08-475-055-8
5	67	94.4	20	1	US-07-755-161A-3
6	67	94.4	20	1	US-07-891-174-3
7	67	94.4	20	1	US-08-204-487-1
8	67	94.4	20	1	US-08-256-771-24
9	67	94.4	20	1	US-08-256-771-25
10	67	94.4	20	1	US-08-381-984-24
11	67	94.4	20	1	US-08-381-984-25
12	67	94.4	22	4	US-09-508-734-4
13	67	94.4	24	4	US-09-508-734-6
14	67	94.4	25	1	US-07-755-161A-10
15	67	94.4	25	1	US-07-891-174-10
16	67	94.4	25	1	US-08-204-487-7
17	67	94.4	29	4	US-09-508-734-8
18	67	94.4	36	1	US-07-755-161A-8
19	67	94.4	36	1	US-07-891-174-8
20	67	94.4	36	1	US-08-256-771-30
21	67	94.4	36	1	US-08-381-984-29
22	67	94.4	47	2	US-08-464-182A-6
23	67	94.4	47	2	US-08-406-271-6
24	67	94.4	50	2	US-08-693-274A-7
25	67	94.4	52	4	US-09-017-043A-3
26	67	94.4	53	2	US-08-464-182A-5
27	67	94.4	53	2	US-08-406-271-5

28	67	94.4	54	2	US-08-464-182A-2	Sequence 2, Appli
29	67	94.4	54	2	US-08-406-271-2	Sequence 2, Appli
30	67	94.4	694	3	US-08-724-586-2	Sequence 2, Appli
31	67	94.4	694	4	US-09-421-632-2	Sequence 2, Appli
32	67	94.4	694	4	US-09-932-190-2	Sequence 2, Appli
33	67	94.4	705	2	US-08-555-640-2	Sequence 2, Appli
34	67	94.4	708	2	US-08-655-640-4	Sequence 4, Appli
35	67	94.4	711	1	US-08-154-019-4	Sequence 4, Appli
36	67	94.4	711	1	US-08-461-333-4	Sequence 4, Appli
37	67	94.4	711	3	US-08-464-167-4	Sequence 4, Appli
38	67	94.4	711	3	US-09-158-313-4	Sequence 4, Appli
39	67	94.4	711	3	US-08-476-798-4	Sequence 4, Appli
40	64	90.1	711	1	US-08-145-681-2	Sequence 2, Appli
41	64	90.1	711	1	US-08-250-308-2	Sequence 2, Appli
42	64	90.1	711	1	US-08-453-703-2	Sequence 2, Appli
43	64	90.1	711	2	US-08-456-106-2	Sequence 2, Appli
44	64	90.1	711	3	US-08-456-108-2	Sequence 2, Appli
45	64	90.1	711	4	US-09-265-577-2	Sequence 2, Appli

#### ALIGNMENTS

RESULT 1  
US-08-204-487-3  
; Sequence 3, Application US/08204487  
; Patent No. 5565425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAOKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGRAKI  
; APPLICANT: DOSAKO, SHUN'ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIAKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESSEE: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FJN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..18  
; OTHER INFORMATION: /note= "HUMAN LACTOFERRIN PEPTIDE"

OTHER INFORMATION: (20-37)"  
US-08-204-487-3

Query Match 94.4%; Score 67; DB 1; Length 18;  
Best Local Similarity 91.7%; Pred. No. 9.9e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWRNMRKVR 12  
Db 1 CFQWRNMRKVR 12

## RESULT 2

US-08-485-948-8  
; Sequence 8, Application US/08485948  
; Patent No. 5855882  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,948  
; FILING DATE:  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/488,217  
; FILING DATE: JUNE 7, 1995  
; APPLICATION NUMBER: 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LF-C1, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
US-08-485-948-8

Query Match 94.4%; Score 67; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 9.9e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWRNMRKVR 12  
Db 1 CFQWRNMRKVR 12

## RESULT 3

US-08-628-380-8  
; Sequence 8, Application US/08628380  
; Patent No. 5891341  
; GENERAL INFORMATION:  
; APPLICANT: LI, YONG MING  
; APPLICANT: VLASSARA, HELEN  
; APPLICANT: CERAMI, ANTHONY  
; TITLE OF INVENTION: AGE-MEDIATED INHIBITION OF ANTIBACTERIAL PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/628,380  
; FILING DATE: April 4, 1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LF-C1, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
US-08-628-380-8

Query Match 94.4%; Score 67; DB 2; Length 18;  
Best Local Similarity 91.7%; Pred. No. 9.9e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWRNMRKVR 12  
Db 1 CFQWRNMRKVR 12

## RESULT 4

US-08-475-055-8  
; Sequence 8, Application US/08475055  
; Patent No. 5962245  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,055
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/485,948
FILING DATE:
APPLICATION NUMBER: 08/488,217
FILING DATE: JUNE 7, 1995
APPLICATION NUMBER: 08/418,642
FILING DATE: APRIL 7, 1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 947-1-008A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 343-1684
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
DESCRIPTION: LP-Cl, 8-25
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
US-08-475-055-8

Query Match 94.4%; Score 67; DB 2; Length 18;
Best Local Similarity 91.7%; Pred No. 9.9e-05;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWRNMEKVR 12
Db 1 CFQWRNMEKVR 12

RESULT 5
US-07-755-161A-3
Sequence 3, Application US/07755161A
Patent No. 5304633
GENERAL INFORMATION:
APPLICANT: Mamoru TOMITA et al.
TITLE OF INVENTION: Antimicrobial Peptides and an
TITLE OF INVENTION: Antimicrobial Agent
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/755,161A
FILING DATE: 19910905
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX: 202-371-8856
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY: modified site
LOCATION: 2
IDENTIFICATION METHOD:
OTHER INFORMATION: /note= "thiol group of
OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with
OTHER INFORMATION: thiol group of Cys residue at location 19"
FEATURE:
NAME/KEY: modified site
LOCATION: 19
IDENTIFICATION METHOD:
OTHER INFORMATION: /note= "thiol group of
OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with
OTHER INFORMATION: thiol group of Cys residue at location 2"
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-07-755-161A-3

Query Match 94.4%; Score 67; DB 1; Length 20;
```



Best Local Similarity 91.7%; Pred. No. 0.00011; Indels 0; Gaps 0;  
Matches 11; Conservative 1; Mismatches 0;  
Qy 1 CFQWRNNRKVR 12  
Db 2 CFQWRNNRKVR 13

RESULT 6  
US-07-891-174-3  
; Sequence 3, Application US/07891174  
; Patent No. 5317084  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: DisplayWrite  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/891,174  
; FILING DATE: 29-MAY-1992  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/755,161  
; FILING DATE: 05-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX: 202-371-8856  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; HYPOTHEetical:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:  
; CELL LINE:  
; ORGANELLE:  
; IMMEDIATE SOURCE:  
; LIBRARY:  
; CLONE:  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT:  
; MAP POSITION:  
; UNITS:  
; FEATURE:  
; NAME/KEY: modified site

LOCATION: 2  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 19"  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 19  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 2"  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-891-174-3

Query Match 94.4%; Score 67; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 0.00011;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CFQWRNNRKVR 12  
Db 2 CFQWRNNRKVR 13

RESULT 7  
US-08-204-487-1  
; Sequence 1, Application US/08204487  
; Patent No. 5565425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAOKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGEAKI  
; APPLICANT: DOSAKO, SHUN ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; TITLE OF INVENTION: INHIBITORS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESSEE: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FJN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..20  
OTHER INFORMATION: /note= "ANTIBACTERIAL PEPTIDE  
DERIVED FROM HUMAN LACTOFERRIN"  
US-08-204-487-1

Query Match 94.4%; Score 67; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 0.00011;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWRNRKVR 12  
| | | | | | | | | |  
Db 2 CFQWRNRKVR 13

RESULT 8  
US-08-256-771-24  
; Sequence 24, Application US/08256771  
; Patent No. 5656591  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
; PRODUCTS THEREWITH  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/256, 771  
; FILING DATE: July 22, 1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "Cys residues are linked by  
; OTHER INFORMATION: disulfide bond"

US-08-256-771-24

Query Match 94.4%; Score 67; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 0.00011;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWRNRKVR 12  
| | | | | | | | | |  
Db 2 CFQWRNRKVR 13

RESULT 9  
US-08-256-771-25  
; Sequence 25, Application US/08256771  
; Patent No. 5656591  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
; PRODUCTS THEREWITH  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/256, 771  
; FILING DATE: July 22, 1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION: /note= "Cys residues are protected to  
; OTHER INFORMATION: prevent disulfide bond"  
US-08-256-771-25

Query Match 94.4%; Score 67; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 0.00011;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWRNRKVR 12  
| | | | | | | | | |  
Db 2 CFQWRNRKVR 13

RESULT 10  
US-08-381-984-24

Sequence 24, Application US/08381984  
Patent No. 5804555  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIOXIDANT  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/381,984  
FILING DATE: April 11, 1995  
CLASSIFICATION: 252  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:

TELEX:  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "cysteine residues at positions 2  
OTHER INFORMATION: and 19 are bonded by disulfide linkage"  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "the specified peptide as well as  
OTHER INFORMATION: peptides including the specified peptide as a fragment thereof"

Query Match 94.4%; Score 67; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 0.00011;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
US-08-381-984-24

QY 1 CFQWQNRMRKVR 12  
DB 2 CFQWQNRMRKVR 13

RESULT 11  
US-08-381-984-25  
Sequence 25, Application US/08381984  
Patent No. 5804555  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIOXIDANT  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack

STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/381,984  
FILING DATE: April 11, 1995  
CLASSIFICATION: 252  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "the specified peptide as well as  
OTHER INFORMATION: peptides including the specified peptide as a fragment thereof"

Query Match 94.4%; Score 67; DB 1; Length 20;  
Best Local Similarity 91.7%; Pred. No. 0.00011;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
US-08-381-984-25

QY 1 CFQWQNRMRKVR 12  
DB 2 CFQWQNRMRKVR 13

RESULT 12  
US-09-508-734-4  
Sequence 4, Application US/09508734  
Patent No. 6423509  
GENERAL INFORMATION:  
APPLICANT: Samyang Genex Corporation  
TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast an  
FILE REFERENCE: PA/SYG/00139  
CURRENT APPLICATION NUMBER: US/09/508,734  
CURRENT FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: PCT/KR99/00373  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: KR1998-29351  
PRIOR FILING DATE: 1998-07-13  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Kopatentin 1.71  
SEQ ID NO 4

```

; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-508-734-4

Query Match          94.4%; Score 67; DB 4; Length 22;
Best Local Similarity 91.7%; Pred. No. 0.00012;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNRNMRKVR 12
Db 2 CFQWQNRNMRKVR 13

RESULT 13
US-09-508-734-6
; Sequence 6, Application US/09508734
; Patent No. 6423509
; GENERAL INFORMATION:
; APPLICANT: Samyang Genex Corporation
; TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and
; FILE REFERENCE: PA/SYG/00139
; CURRENT APPLICATION NUMBER: US/09/508,734
; PRIOR APPLICATION NUMBER: PCT/KR99/00373
; PRIOR FILING DATE: 2000-06-01
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: KR1998-29351
; PRIOR FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: KopatentIn 1.71
; SEQ ID NO 6
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-508-734-6

Query Match          94.4%; Score 67; DB 4; Length 24;
Best Local Similarity 91.7%; Pred. No. 0.00013;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNRNMRKVR 12
Db 3 CFQWQNRNMRKVR 14

RESULT 14
US-07-755-161A-10
; Sequence 10, Application US/07755161A
; Patent No. 5304633
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; TITLE OF INVENTION: Antimicrobial Agent
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/755.161A
; FILING DATE: 19910905
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 4
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 21"
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 21
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 4"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-07-755-161A-10

Query Match          94.4%; Score 67; DB 1; Length 25;
Best Local Similarity 91.7%; Pred. No. 0.00013;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWQNRNMRKVR 12
Db 4 CFQWQNRNMRKVR 15
```

```
RESULT 15
US-07-891-174-10
; Sequence 10, Application US/07891174
; Patent No. 5317084
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; TITLE OF INVENTION: Antimicrobial Agent
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/891,174
; FILING DATE: 29-MAY-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/755,161
; FILING DATE: 05-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 4
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 21"
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 21
```

```
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 4"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-07-891-174-10
;
; Query Match 94.4%; Score 67; DB 1; Length 25;
; Best Local Similarity 91.7%; Pred. No. 0.00013;
; Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 CFQWKENMKVR 12
; DB 4 CFQWQRNMRKVR 15
;
; Search completed: February 21, 2003, 08:04:26
; Job time : 8.93023 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 08:08:15 ; Search time 6.88372 Seconds  
(without alignments)  
54.162 Million cell updates/sec

Title: US-09-743-107b-92  
Perfect score: 71  
Sequence: 1 CFQWRNNRKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 155504 seqs, 31069816 residues

Total number of hits satisfying chosen parameters: 156504

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pdb:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pdb:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pdb:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pdb:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pdb:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pdb:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pdb:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pdb:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pdb:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pdb:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pdb:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pdb:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pdb:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pdb:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67	94.4	15	9	US-09-798-869-2
2	67	94.4	25	9	US-09-798-869-20
3	67	94.4	694	9	US-10-023-096-2
4	59	83.1	15	9	US-09-798-869-6
5	50	70.4	15	9	US-09-798-869-3
6	50	70.4	25	9	US-09-798-869-23
7	42	59.2	15	9	US-09-798-869-7
8	41	57.7	15	9	US-09-798-869-4
9	41	57.7	25	9	US-09-798-869-22
10	39	54.9	333	9	US-09-796-753-26
11	38	53.5	15	9	US-09-798-869-8
12	38	53.5	15	9	US-09-798-869-29
13	38	53.5	15	9	US-09-798-869-30
14	38	53.5	21	10	US-09-864-761-47985
15	38	53.5	489	9	US-09-888-320-2
16	37	52.1	77	10	US-09-864-761-41002
17	37	52.1	86	9	US-09-738-626-5715
18	37	52.1	184	10	US-09-925-301-1248
19	37	52.1	338	9	US-09-978-295A-119

20	37	52.1	338	9	US-09-978-697-119	Sequence 119, App
21	37	52.1	338	9	US-09-978-192A-119	Sequence 119, App
22	37	52.1	338	9	US-09-999-832A-119	Sequence 119, App
23	37	52.1	338	9	US-09-978-189-119	Sequence 119, App
24	37	52.1	553	9	US-09-796-753-14	Sequence 14, Appl
25	37	52.1	553	10	US-09-981-649A-6	Sequence 6, Appl
26	37	52.1	553	10	US-09-981-649A-24	Sequence 24, Appl
27	37	52.1	554	10	US-09-981-649A-30	Sequence 30, Appl
28	37	52.1	554	10	US-09-981-649A-32	Sequence 32, Appl
29	37	52.1	559	10	US-09-981-649A-28	Sequence 28, Appl
30	35	49.3	301	9	US-10-080-960-11	Sequence 11, Appl
31	35	49.3	302	10	US-09-948-078-2	Sequence 2, Appl
32	35	49.3	351	10	US-09-853-625B-16	Sequence 16, Appl
33	35	49.3	1013	9	US-10-028-072-38	Sequence 38, Appl
34	35	49.3	1013	9	US-10-121-049-38	Sequence 38, Appl
35	35	49.3	1013	9	US-10-123-904-38	Sequence 38, Appl
36	35	49.3	1013	9	US-10-140-470-38	Sequence 38, Appl
37	35	49.3	1013	9	US-10-175-746-38	Sequence 38, Appl
38	35	49.3	1013	9	US-10-176-918-38	Sequence 38, Appl
39	35	49.3	1013	9	US-10-176-921-38	Sequence 38, Appl
40	35	49.3	1013	9	US-10-137-865-38	Sequence 38, Appl
41	35	49.3	1013	9	US-10-140-474-38	Sequence 38, Appl
42	34	47.9	62	10	US-09-815-242-12129	Sequence 12129, A
43	34	47.9	62	10	US-09-815-242-13026	Sequence 13026, A
44	34	47.9	509	10	US-09-879-957-194	Sequence 194, App
45	34	47.9	846	9	US-10-051-409-4	Sequence 4, Appl

#### ALIGNMENTS

RESULT 1  
US-09-798-869-2  
; Sequence 2, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PPT  
; ORGANISM: HOMO SAPIENS  
US-09-798-869-2

Query Match 94.4%; Score 67; DB 9; Length 15;  
Best Local Similarity 91.7%; Pred.No. 5.1e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWRNNRKVR 12  
|||:|||||  
DB 3 CFQWRNNRKVR 14

RESULT 2  
US-09-798-869-20  
; Sequence 20, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON

APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 20  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-798-869-20

Query Match 94.4%; Score 67; DB 9; Length 25;  
Best Local Similarity 91.7%; Pred. No. 8.2e-05;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWKNRKVR 12  
| | | | |  
DB 3 CFQWKNRKVR 14

RESULT 3  
US-10-023-096-2  
; Sequence 2, Application US/10023096  
; Patent No. US20020160941A1  
; GENERAL INFORMATION:  
; APPLICANT: Kruzel, Marian L.  
; APPLICANT: Kurecki, Tomasz  
; APPLICANT: Gollnick, Paul D.  
; APPLICANT: Doyle, Darrell J.  
; TITLE OF INVENTION: Cloning, Expression, and Uses of Human  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Jacobson, Price, Holman & Stern  
; STREET: 400 Seventh St. N.W.  
; CITY: Washington D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE:  
; APPLICATION NUMBER: US/10/023,096  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/724,586  
; FILING DATE: 30-SEPT-1996  
; APPLICATION NUMBER: US 08/238,445  
; FILING DATE: 05-MAY-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Player, William E.  
; REGISTRATION NUMBER: 31,409  
; REFERENCE/DOCKET NUMBER: 10505/P58185C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 638-6666  
; TELEFAX: (202) 393-3350  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 694 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-10-023-096-2

Query Match 94.4%; Score 67; DB 9; Length 694;  
Best Local Similarity 91.7%; Pred. No. 0.0017;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWKNRKVR 12  
| | | | |  
DB 22 CFQWKNRKVR 33

RESULT 4  
US-09-798-869-6  
; Sequence 6, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ(RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of Homo sapiens  
; OTHER INFORMATION: sequence)  
US-09-798-869-6

Query Match 83.1%; Score 59; DB 9; Length 15;  
Best Local Similarity 83.3%; Pred. No. 0.001;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWKNRKVR 12  
| | | | |  
DB 3 CFQWKNRKVR 14

RESULT 5  
US-09-798-869-3  
; Sequence 3, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ(RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: CAPRINE  
US-09-798-869-3

Query Match 70.4%; Score 50; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.028;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWKRNMRKV 11  
|:|:|:|:  
Db 3 CYQWQWRMRKL 13

## RESULT 6

US-09-798-869-23  
; Sequence 23, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: CAPRINE  
US-09-798-869-23

Query Match 70.4%; Score 50; DB 9; Length 25;  
Best Local Similarity 63.6%; Pred. No. 0.045;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWKRNMRKV 11  
|:|:|:|:  
Db 3 CYQWQWRMRKL 13

## RESULT 7

US-09-798-869-7  
; Sequence 7, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic peptide (modified form of homo sapiens  
; OTHER INFORMATION: sequence)  
US-09-798-869-7

Query Match 59.2%; Score 42; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.55;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWKRNMRKV 11  
|:|:|:|:  
Db 3 CYQWQWRMRKL 13

## RESULT 8

US-09-798-869-4  
; Sequence 4, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-798-869-4

Query Match 57.7%; Score 41; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.8;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWKRNMRKV 11  
|:|:|:|:  
Db 3 CLRWQWRMRKV 13

## RESULT 9

US-09-798-869-22  
; Sequence 22, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 22  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-798-869-22

Query Match 57.7%; Score 41; DB 9; Length 25;  
Best Local Similarity 54.5%; Pred. No. 1.3;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWKRNMRKV 11  
|:|:|:|:  
Db 3 CLRWQWRMRKV 13



RESULT 10  
US-09-796-753-26  
; Sequence 26, Application US/09796753  
; Publication No. US20030027998A1  
; GENERAL INFORMATION:  
; APPLICANT: McCarthy, Sean A.  
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF  
; FILE REFERENCE: 7853-227-999  
; CURRENT APPLICATION NUMBER: US/09/796,753  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: 09/183,175  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: 09/223,094  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/223,546  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/224,246  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/259,388  
; PRIOR FILING DATE: 1999-02-26  
; PRIOR APPLICATION NUMBER: 60/122,458  
; PRIOR FILING DATE: 1999-03-01  
; PRIOR APPLICATION NUMBER: 09/312,359  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 09/336,536  
; PRIOR FILING DATE: 1999-06-18  
; PRIOR APPLICATION NUMBER: 09/342,687  
; PRIOR FILING DATE: 1999-06-29  
; PRIOR APPLICATION NUMBER: 09/345,464  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: 09/365,164  
; PRIOR FILING DATE: 1999-07-30  
; PRIOR APPLICATION NUMBER: 09/399,723  
; PRIOR FILING DATE: 1999-09-20  
; PRIOR APPLICATION NUMBER: 09/409,634  
; PRIOR FILING DATE: 1999-09-30  
; PRIOR APPLICATION NUMBER: 09/471,179  
; PRIOR FILING DATE: 1999-12-23  
; PRIOR APPLICATION NUMBER: 09/474,071  
; PRIOR FILING DATE: 1999-12-29  
; PRIOR APPLICATION NUMBER: 09/474,072  
; PRIOR FILING DATE: 1999-12-29  
; PRIOR APPLICATION NUMBER: 09/514,010  
; PRIOR FILING DATE: 2000-02-25  
; PRIOR APPLICATION NUMBER: 09/516,745  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 09/572,002  
; PRIOR FILING DATE: 2000-05-14  
; PRIOR APPLICATION NUMBER: 09/597,993  
; PRIOR FILING DATE: 2000-06-19  
; PRIOR APPLICATION NUMBER: 09/599,596  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 09/630,334  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: 09/606,565  
; PRIOR FILING DATE: 2000-06-29  
; PRIOR APPLICATION NUMBER: 09/606,317  
; PRIOR FILING DATE: 2000-06-29  
; PRIOR APPLICATION NUMBER: 09/665,666  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: 09/677,751  
; PRIOR FILING DATE: 2000-09-30  
; NUMBER OF SEQ ID NOS: 162  
; SEQ ID NO 26  
; LENGTH: 333  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-796-753-26

Query Match 54.9%; Score 39; DB 9; Length 333;  
Best Local Similarity 54.5%; Pred. No. 29;  
Matches 6; Conservative 2; Mismatches 3; Indels 3; Gaps 0;

Qy 1 CFQWKNRKV 11  
| : | | | : |  
Db 48 CYQWKRNKGV 58  
| : | | | : |  
RESULT 11  
US-09-798-869-8  
; Sequence 8, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of caprine  
; OTHER INFORMATION: sequence)  
US-09-798-869-8

Query Match 53.5%; Score 38; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 2.4;  
Matches 6; Conservative 2; Mismatches 3; Indels 3; Gaps 0;

Qy 1 CFQWKNRKV 11  
| : | | | : |  
Db 3 CLRQWEMRKV 13  
| : | | | : |

RESULT 12  
US-09-798-869-29  
; Sequence 29, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 29  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: BOVINE  
US-09-798-869-29

Query Match 53.5%; Score 38; DB 9; Length 15;  
Best Local Similarity 45.5%; Pred. No. 2.4;  
Matches 5; Conservative 4; Mismatches 2; Indels 2; Gaps 0;

Qy 1 CFQWKNRKV 11  
| : | | | : |

Db 3 CFWQWRMKKL 13

## RESULT 13

US-09-798-869-30  
; Sequence 30, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVENBJ(RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 30  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: BOVINE  
US-09-798-869-30

Query Match 53.5%; Score 38; DB 9; Length 15;  
Best Local Similarity 45.5%; Pred. No. 2.4;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFWQWRMKKL 11

|||:|:

Db 3 CFWQWRMKKL 13

## RESULT 14

US-09-864-761-47985  
; Sequence 47985, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecmca-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; PRIOR FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 47985  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL096701.14  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6  
; OTHER INFORMATION: EST\_HUMAN HIT: AW294800.1, EVALUE 1.00e-06  
US-09-864-761-47985

Query Match 53.5%; Score 38; DB 10; Length 21;  
Best Local Similarity 83.3%; Pred. No. 3.3;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFWQWR 6

|||:|:

Db 16 CFWQWR 21

## RESULT 15

US-09-888-320-2  
; Sequence 2, Application US/09888320  
; Publication No. US20030013090A1  
; GENERAL INFORMATION:  
; APPLICANT: Barry III, Clifton E.  
; APPLICANT: DeBarber, Andrea E.  
; APPLICANT: Mdluli, Khisimuza  
; APPLICANT: Bekker, Linda-Gail  
; APPLICANT: The Government of the United States of America  
; APPLICANT: as represented by The Secretary of the  
; APPLICANT: Department of Health and Human Services  
; TITLE OF INVENTION: Methods of Diagnosing Multidrug Resistant Tuberculosis  
; FILE REFERENCE: 015280-413100US  
; CURRENT APPLICATION NUMBER: US/09/888,320  
; CURRENT FILING DATE: 2001-06-22  
; PRIOR APPLICATION NUMBER: US 60/214,187  
; PRIOR FILING DATE: 2000-06-26  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 489  
; TYPE: PRT  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: wild-type Etad monooxygenase (Rv3854c, EthA)  
US-09-888-320-2

Query Match 53.5%; Score 38; DB 9; Length 489;  
Best Local Similarity 54.5%; Pred. No. 59;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFWQWRMKKL 11

|||:|:

Db 253 CQWPRRVRKRM 263

Search completed: February 21, 2003, 08:11:56  
Job time : 6.88372 secs

---

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:48:01 ; Search time 10.6047 Seconds  
(without alignments)  
108.784 Million cell updates/sec

Title: US-09-743-107B-92

Perfect score: 71

Sequence: 1 CFQWKNRKR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 73:\*

1: Pirl:\*

2: Pirl2:\*

3: Pirl3:\*

4: Pirl4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67	94.4	711	1 TFHUL	lactotransferrin precursor
2	50	70.4	708	2 JC2323	lactoferrin - goat
3	47	66.2	33	2 S52107	lactoferrin - sheep
4	45	63.4	275	2 T22597	hypothetical prote
5	41	57.7	206	2 H97451	pyridoxamine 5'-ph
6	41	57.7	206	2 AB2670	probable pyridoxam
7	41	57.7	208	2 AG3441	hypothetical prote
8	41	57.7	511	2 AB0858	hypothetical prote
9	41	57.7	707	1 A28438	lactoferrin precu
10	40	56.3	584	2 C94325	hypothetical prote
11	40	56.3	4568	2 T08030	dyein beta heavy
12	39	54.9	447	2 T18633	hypothetical prote
13	39	54.9	759	2 G86506	hypothetical prote
14	39	54.9	759	2 G72115	hypothetical prote
15	38	53.5	60	2 A48336	ribosomal protein
16	38	53.5	62	2 AH1301	ribosomal protein
17	38	53.5	62	2 AH1673	ribosomal protein
18	38	53.5	267	2 S77802	hypothetical prote
19	38	53.5	298	2 A23466	hypothetical prote
20	38	53.5	393	2 E64639	3-deoxy-manno-octu
21	38	53.5	489	2 C70655	probable monooxyge
22	38	53.5	583	2 T01470	diphosphate-fructo
23	37	52.1	205	2 E90094	26S proteasome SU
24	37	52.1	214	2 S07989	vif protein - simi
25	37	52.1	274	2 E50950	apolipoprotein B-1
26	37	52.1	357	2 T22879	hypothetical prote
27	37	52.1	513	2 E86156	T14P4.7 protein -
28	37	52.1	518	2 B84514	probable cytochrom
29	37	52.1	536	2 T24218	hypothetical prote

## RESULT 1

TFHUL

lactotransferrin precursor [validated] - human

N/Alternate names: lactoferrin

C/Species: Homo sapiens (man)

C/Date: 31-Mar-1992 #sequence revision 21-Nov-1997 #text change 08-Dec-2000

C/Accession: G01394; S11228; A45401; S10324; S15853; S20841; S07160; A61169; A31000; S

R/Chco. Y.

submitted to the EMBL Data Library, March 1994

A/Reference number: G06820

A/Accession: G01394

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-711 <CHO>

A/Cross-references: EMBL:U07643; NID:G467236; PIDN:AAB60324.1; PID:G467237

R/Rey, M.W.; Woloshuk, S.L.; deBoer, H.A.; Pieper, F.R.

Nucleic Acids Res. 18, 5288, 1990

A/Title: Complete nucleotide sequence of human mammary gland lactoferrin.

A/Reference number: S11228; MUID:90384839; PMID:2402455

A/Accession: S11228

A/Molecule type: mRNA

A/Residues: 1-148; 'T', 150-422; 'C', 424-711 <REY>

A/Cross-references: EMBL:X53961; NID:G34415; PIDN:CAA37914.1; PID:G34416

R/Teng, C.T.; Liu, Y.; Yang, N.; Walmer, D.; Panella, T.

Mol. Endocrinol. 6, 1969-1981, 1992

A/Title: Differential molecular mechanism of the estrogen action that regulates lactofe

A/Reference number: A45401; MUID:93125571; PMID:1480183

A/Accession: A45401

A/Molecule type: DNA

A/Residues: 1-15 <TEN>

A/Cross-references: GB:S52659; NID:G263311; PIDN:AAB24877.1; PID:G263312

A/Experimental source: placenta

A/Note: sequence extracted from NCBI backbone (NCBIP:122202)

R/Powell, M.J.; Ogden, J.E.

Nucleic Acids Res. 18, 4013, 1990

A/Title: Nucleotide sequence of human lactoferrin cDNA.

A/Reference number: S10324; MUID:90326549; PMID:2374734

A/Accession: S10324

A/Molecule type: mRNA

A/Residues: 3-711 <POW>

A/Cross-references: EMBL:X52941; NID:G34411; PIDN:CAA37116.1; PID:G34412

R/Stowell, K.M.; Rado, T.A.; Funk, W.D.; Tweedie, J.W.

Biochem. J. 276, 349-355, 1991

A/Title: Expression of cloned human lactoferrin in baby-hamster kidney cells.

A/Reference number: S15853; MUID:91264786; PMID:2049066

A/Accession: S15853

A/Status: nucleic acid sequence not shown; not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 20-31 <STI>

A/Accession: S20841

A/Molecule type: protein

A/Residues: 20-28; 'X', 30-31 <ST2>

R;Rado, T.A.; Wei, X.; Benz Jr., E.J.  
 Blood 70, 989-993, 1987  
 A;Title: Isolation of lactoferrin cDNA from a human myeloid library and expression of mRNA  
 A;Reference number: S07160; MUID:88001031; PMID:3477300  
 A;Accession: S07160  
 A;Molecule type: mRNA  
 A;Residues: 436-487, 'A', 489-711 <RAD>  
 A;Cross-references: EMBL:M18642; NID:G186815; PIDN:AAA86665.1; PID:G386855  
 R;Panella, T.J.; Liu, Y.; Huang, A.T.; Teng, C.T.  
 Cancer Res. 51, 3037-3043, 1991  
 A;Title: Polymorphism and altered methylation of the lactoferrin gene in normal leukocytes  
 A;Reference number: A61169; MUID:91235214; PMID:1674448  
 A;Accession: A61169  
 A;Status: not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 3-701, 'SWKPNV' <PAN>  
 A;Experimental source: normal breast tissue  
 R;Metz-Boutigue, M.H.; Jolles, J.; Mazurier, J.; Schoentgen, F.; Legrand, D.; Spik, G.;  
 Eur. J. Biochem. 145, 659-666, 1984  
 A;Title: Human lactotransferrin: amino acid sequence and structural comparisons with other  
 A;Reference number: A31000; MUID:85076667; PMID:6510420  
 A;Accession: A31000  
 A;Molecule type: protein  
 A;Residues: 20-140, 142-169, 171-203, 'L', 205, 'K', 207-208, 'K', 210-385, 'Q', 387-391, 'W', 393-4  
 A;Note: this is the final paper in a series  
 R;Houen, G.; Hoegdall, E.V.; Barkholt, V.; Norskov, L.  
 Eur. J. Biochem. 241, 303-308, 1996  
 A;Title: Lactoferrin: similarity to diamine oxidase and purification by aminohexyl affinity  
 A;Reference number: S74119; MUID:97054624; PMID:8898921  
 A;Accession: S74119  
 A;Molecule type: protein  
 A;Residues: 'G', 23-24, 'R', 26-27, 'XX', 30-32 <HOU>  
 A;Experimental source: neutrophil granulocytes  
 C;Genetics:  
 A;Gene: GDB:LNFB  
 A;Cross-references: GDB:119368; OMIM:150210  
 A;Map position: 3q21-3q23  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein; iron binding; milk  
 F;1-19/Domain: signal sequence #status predicted <SIG>  
 F;20-711/Product: lactotransferrin #status experimental <MAT>  
 F;21-356/Domain: transferrin repeat homology <TRH1>  
 F;360-699/Domain: transferrin repeat homology <TRH2>  
 F;29-65,39-56,135-218,177-193,190-201,251-265,503-697,595-609/Disulfide bonds: #status e  
 F;157,498/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F;362-400,378-391,425-706,447-669,479-554,513-527,524-537,647-652/Disulfide bonds: #statu

Query Match 94.4%; Score 67; DB 1; Length 711;  
 Best Local Similarity 91.7%; Pred. No. 0.00075;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWKNRKVR 12  
 DB 39 CFQWQNRKVR 50

RESULT 2  
 JC2323  
 lactoferrin - goat  
 C;Species: Capra aegagrus hircus (domestic goat)  
 C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 07-May-1999  
 C;Accession: JC2323  
 R;Le Provost, F.; Nocart, M.; Querin, G.; Martin, P.  
 Biochem. Biophys. Res. Commun. 203, 1324-1332, 1994  
 A;Title: Characterization of the goat lactoferrin cDNA: assignment of the relevant locus  
 A;Reference number: JC2323; MUID:94380047; PMID:8093048  
 A;Accession: JC2323  
 A;Molecule type: mRNA  
 A;Residues: 1-708 <LEP>  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein  
 F;559-686/Domain: transferrin repeat homology <TRH2>  
 F;1252,300,387,495,564/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 70.4%; Score 50; DB 2; Length 708;  
 Best Local Similarity 63.6%; Pred. No. 0.072;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWKNRKVR 11  
 DB 38 CFQWQNRKVR 48

RESULT 3  
 SS2107  
 lactoferrin - sheep (fragment)  
 C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C;Date: 14-Jul-1995 #sequence\_revision 05-Dec-1997 #text\_change 05-Dec-1997  
 C;Accession: SS2107  
 R;Qian, Z.Y.; Jolles, P.; Migliore-Samour, D.; Fiat, A.M.  
 Biochim. Biophys. Acta 1243, 25-32, 1995  
 A;Title: Isolation and characterization of sheep lactoferrin, an inhibitor of platelet  
 A;Reference number: SS2107; MUID:95127729; PMID:7827104  
 A;Accession: SS2107  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-33 <QIA>  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication

Query Match 66.2%; Score 47; DB 2; Length 33;  
 Best Local Similarity 54.5%; Pred. No. 0.12;  
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWKNRKVR 11  
 DB 19 CFQWQNRKVR 29

RESULT 4  
 T22597  
 hypothetical protein F53H4.4 - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
 C;Accession: T22597  
 R;Dobson, R.  
 Submitted to the EMBL Data Library, October 1996  
 A;Reference number: Z19587  
 A;Accession: T22597  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-275 <WIL>  
 A;Cross-references: EMBL:Z81089; PIDN:CAB03137.1; GSPDB:GN00028; CESP:F53H4.4  
 A;Experimental source: clone F53H4  
 C;Genetics:  
 A;Gene: CESP:F53H4.4  
 A;Map position: X  
 A;Introns: 67/1; 153/1  
 C;Superfamily: Caenorhabditis elegans hypothetical protein F53H4.4

Query Match 63.4%; Score 45; DB 2; Length 275;  
 Best Local Similarity 72.7%; Pred. No. 2.2;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQWKISMRKVR 12  
 DB 262 FQWKISMRKTR 272

RESULT 5  
 H97451  
 Pyridoxamine 5'-phosphate oxidase (AP179611) [imported] - Agrobacterium tumefaciens (s  
 C;Species: Agrobacterium tumefaciens  
 C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002  
 C;Accession: H97451  
 R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldma

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001  
 A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens* strain C58, 2001  
 A;Accession: H97451  
 A;Accession: H97451  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-206 <KUR>  
 A;Cross-references: GB:AB007869; PIDN:AAK8569.1; PID:gl15155733; GSPDB:GN00169  
 C;Genetics:  
 A;Gene: AGR C 1381  
 A;Map position: circular chromosome  
 C;Superfamily: pyridoxamine-phosphate oxidase

Query Match 57.7%; Score 41; DB 2; Length 206;  
 Best Local Similarity 58.3%; Pred. No. 8.2;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFOWKRNMRKVR 12  
 |||||  
 Db 88 CFHWKSLRRQVR 99  
 |||||

RESULT 6  
 AB2670  
 A;Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
 A;Reference number: AB2577; PMID:11743193  
 A;Accession: AB2670  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-206 <KUR>  
 A;Cross-references: GB:AB008688; PIDN:AAU41776.1; PID:gl7739129; GSPDB:GN00186  
 A;Experimental source: strain C58 (Dupont)  
 C;Genetics:  
 A;Gene: pdxH  
 A;Map position: circular chromosome  
 C;Superfamily: pyridoxamine-phosphate oxidase

Query Match 57.7%; Score 41; DB 2; Length 206;  
 Best Local Similarity 58.3%; Pred. No. 8.2;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFOWKRNMRKVR 12  
 |||||  
 Db 88 CFHWKSLRRQVR 99  
 |||||

RESULT 7  
 AG3441  
 A;Title: The Genome of the facultative intracellular pathogen *Brucella melitensis* (strain 16M)  
 A;Reference number: AB3252; PMID:11756688  
 A;Accession: AG3441  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-208 <KUR>  
 A;Cross-references: GB:AB008917; PIDN:AAL52698.1; PID:gl7983525; GSPDB:GN00190

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001  
 A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens* strain C58, 2001  
 A;Accession: H97451  
 A;Accession: H97451  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-206 <KUR>  
 A;Cross-references: GB:AB007869; PIDN:AAK8569.1; PID:gl15155733; GSPDB:GN00169  
 C;Genetics:  
 A;Gene: AGR C 1381  
 A;Map position: circular chromosome  
 C;Superfamily: pyridoxamine-phosphate oxidase

Query Match 57.7%; Score 41; DB 2; Length 206;  
 Best Local Similarity 58.3%; Pred. No. 8.2;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFOWKRNMRKVR 12  
 |||||  
 Db 88 CFHWKSLRRQVR 99  
 |||||

RESULT 6  
 AB2670  
 A;Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
 A;Reference number: AB2577; PMID:11743193  
 A;Accession: AB2670  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-206 <KUR>  
 A;Cross-references: GB:AB008688; PIDN:AAU41776.1; PID:gl7739129; GSPDB:GN00186  
 A;Experimental source: strain C58 (Dupont)  
 C;Genetics:  
 A;Gene: pdxH  
 A;Map position: circular chromosome  
 C;Superfamily: pyridoxamine-phosphate oxidase

Query Match 57.7%; Score 41; DB 2; Length 206;  
 Best Local Similarity 58.3%; Pred. No. 8.2;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFOWKRNMRKVR 12  
 |||||  
 Db 88 CFHWKSLRRQVR 99  
 |||||

RESULT 7  
 AG3441  
 A;Title: The Genome of the facultative intracellular pathogen *Brucella melitensis* (strain 16M)  
 A;Reference number: AB3252; PMID:11756688  
 A;Accession: AG3441  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-208 <KUR>  
 A;Cross-references: GB:AB008917; PIDN:AAL52698.1; PID:gl7983525; GSPDB:GN00190

A;Experimental source: strain 16M  
 C;Genetics:  
 A;Gene: BME11517  
 A;Map position: I  
 C;Superfamily: pyridoxamine-phosphate oxidase  
 C;Keywords: oxidoreductase

Query Match 57.7%; Score 41; DB 2; Length 208;  
 Best Local Similarity 58.3%; Pred. No. 8.3;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFOWKRNMRKVR 12  
 |||||  
 Db 90 CFHWKSLRRQVR 101  
 |||||

RESULT 8  
 AB0958  
 A;Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhimurium strain 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 09-Nov-2001  
 A;Accession: AB0958  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-511 <PAR>  
 A;Cross-references: GB:AL513392; PIDN:CAD06049.1; PID:gl6504016; GSPDB:GN00176  
 C;Genetics:  
 A;Gene: STY3070

Query Match 57.7%; Score 41; DB 2; Length 511;  
 Best Local Similarity 58.3%; Pred. No. 20;  
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CFOWKRNMRKVR 12  
 |||||  
 Db 350 CFAMDNRKAKVR 361  
 |||||

RESULT 9  
 A28438  
 A;Title: Lactoferrin precursor - mouse  
 N;Alternate names: lactotransferrin  
 C;Species: Mus musculus (house mouse)  
 C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C;Accession: A28438; A41205  
 R;Pentecost, B.T.; Teng, C.T.  
 J. Biol. Chem. 262, 10134-10139, 1987  
 A;Title: Lactotransferrin is the major estrogen inducible protein of mouse uterine secretory epithelial cells  
 A;Reference number: A92596; PMID:87280033; PMID:3611056  
 A;Accession: A28438  
 A;Molecule type: mRNA  
 A;Residues: 3-707 <PEN>  
 A;Cross-references: EMBL:J03298  
 R;Liu, Y.; Teng, C.T.  
 J. Biol. Chem. 266, 21880-21885, 1991  
 A;Title: Characterization of estrogen-responsive mouse lactoferrin promoter.  
 A;Reference number: A41205; PMID:92042099; PMID:1939912  
 A;Accession: A41205  
 A;Molecule type: DNA  
 A;Residues: 1-15 <LIU>  
 A;Cross-references: GB:M74778  
 C;Superfamily: transferrin; transferrin repeat homology  
 C;Keywords: duplication; glycoprotein  
 F;1-19/Domain: signal sequence #status predicted <SIG>

F:20-707/Product: lactotransferrin #status predicted <MAT>  
 F:358-695/Domain: transferrin repeat homology <TRH2>  
 F:494/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 57.7%; Score 41; DB 1; Length 707;  
 Best Local Similarity 54.5%; Pred. No. 28;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFOWKRNMRKV 11  
 |||:|:|:|:  
 DB 37 CLFQWQEMRKV 47

## RESULT 10

C84325  
 hypothetical protein Vngl732c [imported] - Halobacterium sp. NRC-1  
 C:Species: Halobacterium sp. NRC-1  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: C84325  
 R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laeky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo  
 Jung, K.H.; Alam, M.; Freitas, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
 A:Title: Genome sequence of Halobacterium species NRC-1.  
 A:Reference number: A84160; MUID:20504483; PMID:11016950  
 A:Accession: C84325  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-584 <STO>  
 A:Cross-references: GB:AE004437; NID:gi0581192; PIDN:AAG19967.1; GSPDB:GN00138  
 C:Genetics:  
 A:Gene: VNG1732C

Query Match 56.3%; Score 40; DB 2; Length 584;  
 Best Local Similarity 41.7%; Pred. No. 34;  
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFOWKRNMRKV 12  
 |||:|:|:|:  
 DB 445 CFTWRKDMERK 456

## RESULT 11

T08030  
 dynein beta heavy chain - Chlamydomonas reinhardtii  
 C:Species: Chlamydomonas reinhardtii  
 C:Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 02-Feb-2001  
 C:Accession: T08030  
 R:Mitchell, D.R.; Brown, K.S.  
 J. Cell Sci. 107, 635-644, 1994  
 A:Title: Sequence analysis of the Chlamydomonas alpha and beta dynein heavy chain genes.  
 A:Reference number: Z16302; MUID:94274778; PMID:8006077  
 A:Accession: T08030  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-4568 <MIT>  
 A:Cross-references: EMBL:U02963; NID:g409965; PIDN:AAA19956.1; PID:9514215  
 A:Experimental source: strain 21gr  
 C:Genetics:  
 A:Gene: ODA4

A:Map position: IX  
 A:Introns: 49/2; 74/3; 105/3; 159/1; 225/3; 358/3; 514/3; 729/2; 919/3; 1004/3; 1115/3;  
 3334/3; 3686/3; 3892/3; 4240/3  
 C:Superfamily: dynein heavy chain, ciliary  
 C:Keywords: nucleotide binding; P-loop  
 F:1919-1926/Region: nucleotide-binding motif A (P-loop)  
 F:2202-2209/Region: nucleotide-binding motif A (P-loop)  
 F:2530-2537/Region: nucleotide-binding motif A (P-loop)

Query Match 56.3%; Score 40; DB 2; Length 4568;  
 Best Local Similarity 41.7%; Pred. No. 2.5e+02;  
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFOWKRNMRKV 12  
 |||:|:|:|:  
 DB 1852 CFOWSQRLRYIQ 1863

## RESULT 12

TI8633  
 hypothetical protein M18.8 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: TI8633; T23799  
 R:Sims, M.  
 submitted to the EMBL Data Library, February 1996  
 A:Reference number: Z18999  
 A:Accession: TI8633  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-447 <WIL>  
 A:Cross-references: EMBL:Z69634; PIDN:CAA93457.1; GSPDB:GN000022; CESP:M18.8  
 A:Experimental source: clone B0001  
 R:Steward, C.

submitted to the EMBL Data Library, January 1996

A:Reference number: Z19800  
 A:Accession: T23799  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-447 <W12>  
 A:Cross-references: EMBL:Z68507; PIDN:CAA92831.1; GSPDB:GN000022; CESP:M18.8  
 A:Experimental source: clone M18  
 C:Genetics:  
 A:Gene: CESP:M18.8  
 A:Map position: 4  
 A:Introns: 41/3; 137/1; 326/3; 434/2

Query Match 54.9%; Score 39; DB 2; Length 447;  
 Best Local Similarity 75.0%; Pred. No. 39;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 WKRNMRKV 11  
 |||:|:|:|:  
 DB 291 WKRNLRV 298

## RESULT 13

G86506  
 hypothetical protein CFJ0126 [imported] - Chlamydomophila pneumoniae (strain J138)  
 C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Mar-2001  
 C:Accession: G86506  
 R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;  
 Nucleic Acids Res. 28, 2311-2314, 2000  
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.  
 A:Reference number: A86491; MUID:20330349; PMID:10871362  
 A:Accession: G86506  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-759 <STO>  
 A:Cross-references: GB:BA000008; NID:98978500; PIDN:BAA98337.1; GSPDB:GN00142  
 A:Experimental source: strain J138  
 C:Genetics:  
 A:Gene: CFJ0126

Query Match 54.9%; Score 39; DB 2; Length 759;  
 Best Local Similarity 66.7%; Pred. No. 66;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 QWKRNMRKV 11  
 |||:|:|:|:  
 DB 488 QWKKNLRDV 496

## RESULT 14

G72115  
 hypothetical protein CP0646 [imported] - Chlamydomydia pneumoniae (strains CWL029 and AR  
 C:Species: Chlamydomydia pneumoniae, Chlamydia pneumoniae  
 C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-May-2000  
 C:Accession: G72115; F81554  
 R:Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Pan, J.; Olinger, L.; Grimwood, J.;  
 Nature Genet. 21, 385-389, 1999  
 A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
 A:Reference number: A72000; MUID:99206606; PMID:10192388  
 A:Accession: G72115  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-759 <ARN>  
 A:CROSS-references: GB:AE001599; GB:AE001363; MID:g4376387; PIDN:AAF18279.1; PID:g437639  
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,  
 C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
 Nucleic Acids Res. 28, 1397-1406, 2000  
 A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
 A:Reference number: A81500; MUID:20150255; PMID:10684935  
 A:Accession: F81554  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-759 <REA>  
 A:CROSS-references: GB:AE002222; GB:AE002161; MID:g7189553; PIDN:AAF38461.1; PID:g718956  
 R:Experimental source: strain AR39, HL cells  
 C:Genetics:  
 A:Gene: CP0126; CP0646

Query Match 54.9%; Score 39; DB 2; Length 759;  
 Best Local Similarity 66.7%; Pred.No. 66;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 QWKNMKV 11  
 |||:|:|  
 DB 488 QWKNLRDV 496

RESULT 15  
 A48396  
 ribosomal protein L28 - Bacillus stearothermophilus  
 C:Species: Bacillus stearothermophilus  
 C:Date: 19-Nov-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Aug-1998  
 C:Accession: A48396  
 R:Kruft, V.; Kapp, U.; Wittmann-Liebold, B.  
 Biochimie 73, 855-860, 1991  
 A:Title: Characterization and primary structure of proteins L28, L33 and L34 from Bacill  
 A:Reference number: A48396; MUID:92075758; PMID:1742360  
 A:Accession: A48396  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-60 <KRU>  
 A:Note: sequence extracted from NCBI backbone (NCBIP:69662)  
 C:Superfamily: Escherichia coli ribosomal protein L28

Query Match 53.5%; Score 38; DB 2; Length 60;  
 Best Local Similarity 66.7%; Pred.No. 8.2;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 WKNMKV 12  
 |||:|:|  
 DB 27 WKNLQKV 35

Search completed: February 21, 2003, 08:02:46  
 Job time : 10.6047 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:28:06 ; Search time 5.2093 Seconds  
(without alignments)  
95.544 Million cell updates/sec

Title: US-09-743-107b-92

Perfect score: 71

Sequence: 1 CFQWKNNRKYR 12

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67	94.4	711	1 TRFL_HUMAN	P02788 homo sapien
2	50	70.4	708	1 TRFL_CAMDR	Q9tum0 camelus dro
3	50	70.4	708	1 TRFL_CAPHI	Q29477 capra hircu
4	42	59.2	695	1 TRFL_HORSE	Q77811 equus cabal
5	41	57.7	62	1 RL28_THETN	Q8r9u1 thermoanaer
6	41	57.7	707	1 TRFL_MOUSE	P08071 mus musculu
7	40	56.3	4568	1 DYHB_CHLRE	Q39565 chlamydomon
8	39	54.9	292	1 NLA_DROME	Q9xz18 drosophila
9	38	53.5	60	1 RL28_BACST	P23374 bacillus st
10	38	53.5	62	1 RL28_LISMO	Q92aj2 listeria mo
11	38	53.5	146	1 RPOB_LIBAF	P41187 liberibacte
12	38	53.5	267	1 Y125_MYCCA	P53661 mycoplasma
13	37	52.1	214	1 VIF_SIVS4	P12505 simian immu
14	37	52.1	783	1 YNR2_CAEEL	Q21588 caenorhabdi
15	36	50.7	365	1 IAK3_HUMAN	P30453 homo sapien
16	36	50.7	455	1 YKYL_CAEEL	Q19910 caenorhabdi
17	36	50.7	485	1 GLGA_BACST	Q08328 bacillus st
18	36	50.7	502	1 C932_SOYBN	Q42799 glycine max
19	36	50.7	509	1 C931_SOYBN	Q42798 glycine max
20	36	50.7	510	1 C933_SOYBN	Q81973 glycine max
21	36	50.7	528	1 CAX2_ARATH	Q38798 arabidopsis
22	36	50.7	530	1 CAX1_ARATH	P29402 arabidopsis
23	36	50.7	728	1 KDGI_ARATH	Q39017 arabidopsis
24	36	50.7	749	1 VP4_ROTGA	Q04916 rotavirus (
25	36	50.7	1135	1 PHYC_SORBI	P93528 sorghum bic
26	35	49.3	151	1 SYB2_RHIME	Q923q1 rhizobium m
27	35	49.3	215	1 VIF_HV2SB	P12452 human immu
28	35	49.3	215	1 VIF_HV2ST	P20878 human immu
29	35	49.3	343	1 SP2D_BACSU	P07372 bacillus su
30	35	49.3	344	1 SP2D_BACAM	P13251 bacillus am
31	35	49.3	351	1 NOV_CHICK	P28686 gallus gall
32	35	49.3	358	1 DDL_ENTHR	Q47827 enterococu
33	35	49.3	398	1 YK13_CAEEL	P34337 caenorhabdi

#### ALIGNMENTS

##### RESULT 1

ID	TRFL_HUMAN	STANDARD	PRT	711 AA
AC	P02788; Q16780; Q16785; Q16786; Q16789; Q00756; Q9H123; Q96K24;			
AC	Q96K25;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Lactotransferrin precursor (Lactoferrin) [Contains: Lactoferrin A;			
DE	Lactoferrin B; Lactoferrin C].			
GN	LfR OR LF.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RX	MEDLINE=90384839; PubMed=2402455;			
RA	Rey M.W., Woloshuk S.L., de Boer H.A., Pieper F.R.;			
RT	"Complete nucleotide sequence of human mammary gland lactoferrin.";			
RL	Nucleic Acids Res. 18:5288-5288(1990).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Cho Y.Y.;			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Connely O.M.;			
RT	Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Liang Q., Jimenez-Flores R., Richardson T.;			
RT	"Molecular cloning and sequence analysis of human lactoferrin.";			
RL	Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Bone marrow;			
RA	Wei X., Han J., Rado T.A.;			
RT	"Human neutrophil lactoferrin coding and 5' flanking region DNA			
RT	sequences.";			
RL	Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Cheng H., Chen X., Huan L.;			
RT	"cDNA cloning and sequence analysis of human lactoferrin.";			
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RN	[7]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Prostate;			
RA	Strausberg R.;			
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.			
RN	[8]			

Q9vfn2 drosophila  
Q9pj99 chlamydia m  
Q9ute7 schizosacch  
P35649 eikenella c  
P58306 thermoplaem  
Q9ulc6 homo sapien  
Q99up4 staphylococ  
P29328 oviss aries  
Q10353 schizosacch  
P43488 mus musculu  
P19506 simian immu  
Q55185 synechocyst

RP SEQUENCE OF 3-711 FROM N.A.  
RC TISSUE-Mammary Gland;  
RX MEDLINE=90326549; PubMed=2374734;  
RA Powell M.J., Ogden J.E.;  
RT "Nucleotide sequence of human lactoferrin cDNA.";  
RL Nucleic Acids Res. 18:4013-4013(1990).  
RN [9]  
RP SEQUENCE OF 20-711.  
RX MEDLINE=8507667; PubMed=6510420;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Schoentgen F.,  
RG Legrand D., Spik G., Montreuil J., Jolles P.;  
RT "Human lactotransferrin: amino acid sequence and structural  
RL comparisons with other transferrins.";  
RN Eur. J. Biochem. 145:659-666(1984).  
RP PRELIMINARY SEQUENCE OF 20-73; 134-171; 257-278 AND 347-530.  
RX MEDLINE=82046817; PubMed=6794640;  
RA Metz-Boutigue M.-H., Mazurier J., Jolles J., Spik G., Montreuil J.,  
RG Jolles P.;  
RT "The present state of the human lactotransferrin sequence. Study and  
RL alignment of the cyanogen bromide fragments and characterization of  
RN N- and C-terminal domains.";  
RX Biochim. Biophys. Acta 670:243-254(1981).  
RN [11]  
RP SEQUENCE OF 609-711.  
RX MEDLINE=82262043; PubMed=7049727;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,  
RG Jolles P.;  
RT "An 88 amino acid long C-terminal sequence of human  
RL lactotransferrin.";  
RN FEBS Lett. 142:107-110(1982).  
RP SEQUENCE OF 436-711 FROM N.A.  
RX MEDLINE=88001031; PubMed=3477300;  
RA Rado T.A., Wei X., Benz E.J. Jr.;  
RT "Isolation of lactoferrin cDNA from a human myeloid library and  
RL expression of mRNA during normal and leukemic myelopoiesis.";  
RN Blood 70:989-993(1987).  
RP SEQUENCE OF 237-711 FROM N.A.  
RX McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,  
RA Nhar M., Parnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,  
RG Chen C.N., Evans C., Kaplan N., Greco T., Touchman J., Muzny D.,  
RA Dragan Y., Giacalone J., Pae A., Powell E., Solinsky K.A., Desilva U.,  
RG Diaz-Perez S., Zhou X., Yu Y., Watanabe M., Doggett N., Garcia D.,  
RA Sacripanti J.L.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
RN [14]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND REVISIONS.  
RX MEDLINE=90064528; PubMed=2585506;  
RA Anderson B.F., Baker H.M., Norris G.E., Rice D.W., Baker E.N.;  
RT "Structure of human lactoferrin: crystallographic structure analysis  
RL and refinement at 2.8-A resolution.";  
RN J. Mol. Biol. 209:711-734(1989).  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX Haridas M., Anderson B.F., Baker E.N.;  
RT "Structure of human diferric lactoferrin refined at 2.2-A  
RL resolution.";  
RN Acta Crystallogr. D 51:629-646(1995).  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 20-353.  
RX MEDLINE=97156796; PubMed=9003186;  
RA Nicholson H., Anderson B.F., Bland T., Shewry S.C., Tweedie J.W.,  
RG Baker E.N.;  
RT "Mutagenesis of the histidine ligand in human lactoferrin: iron  
RL binding properties and crystal structure of the histidine-  
RN 253--methionine mutant.";  
RX Biochemistry 36:341-346(1997).  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX MEDLINE=99190892; PubMed=10089347;  
RA Sun X.L., Baker H.M., Shewry S.C., Jameson G.B., Baker E.N.;  
RT "Structure of recombinant human lactoferrin expressed in Aspergillus  
RL awamori.";  
RN Acta Crystallogr. D 55:403-407(1999).  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=99192677; PubMed=10089508;  
RA Jameson G.B., Anderson B.F., Norris G.E., Thomas D.H., Baker E.N.;  
RT "Structure of human apolactoferrin at 2.0-A resolution. Refinement  
RL and analysis of ligand-induced conformational change.";  
RN Acta Crystallogr. D 54:1319-1335(1998).  
RP CHARACTERIZATION OF LACTOFERROXINS.  
RX MEDLINE=91166929; PubMed=1369293;  
RA Tani F., Iio K., Chiba H., Yoshikawa M.;  
RT "Isolation and characterization of opioid antagonist peptides derived  
RL from human lactoferrin.";  
RN Agric. Biol. Chem. 54:1803-1810(1990).  
RP VARIANTS THR-30 AND ARG-48.  
RX PubMed=9873069;  
RA Klintworth G.K., Sommer J.R., Obrian G., Han L., Ahmed M.N.,  
RG Qumsiyeh M.B., Lin P.-Y., Basti S., Reddy M.K., Kanai A., Hotta Y.,  
RA Sugar J., Kumaramanickavel G., Munier F., Schorderet D.F.,  
RG El Matiri L., Iwata F., Kaiser-Kupfer M., Nagata M., Nakayasu K.,  
RA Hejtmancik J.F., Teng C.T.;  
RT "Familial subepithelial corneal amyloidosis (gelatinous drop-like  
RL corneal dystrophy): exclusion of linkage to lactoferrin gene.";  
RN Mol. Vision 4:31-32(1998).  
RP FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
CC OF AN ANION, USUALLY BICARBONATE.  
CC -1- FUNCTION: LACTOFERROXINS A, B AND C HAVE OPIOID ANTAGONIST  
CC ACTIVITY. LACTOFERROXIN A SHOWS PREFERENCE FOR MU-RECEPTORS, WHILE  
CC LACTOFERROXIN B AND C HAVE SOMEWHAT HIGHER DEGREES OF PREFERENCE  
CC FOR KAPPA-RECEPTORS THAN FOR MU-RECEPTORS.  
CC -1- SUBUNIT: MONOMER  
CC -1- SUBCELLULAR LOCATION: Secreted  
CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X53961; CAA37914.1; -  
CC EMBL; U07643; AAB60324.1; -  
CC EMBL; M93150; AAA36159.1; -  
CC EMBL; M83202; AAA59511.1; -  
CC EMBL; M83205; AAA58656.1; -  
CC EMBL; M18642; AAA86665.1; -  
CC EMBL; AF332168; AAG48753.1; -  
CC EMBL; BC015822; AAH15822.1; -  
CC EMBL; BC015823; AAH15823.1; -  
CC EMBL; M73700; AAA59479.1; -  
CC EMBL; X52941; CAA37116.1; -  
CC EMBL; U95626; AAB57795.1; -  
CC PIR; S11228; TFHUL.  
CC PDB; 1LCT; 31-AUG-94.  
CC PDB; 1LCF; 31-OCT-93.  
CC PDB; 1LFG; 31-JUL-94.  
CC PDB; 1LFH; 31-OCT-93.  
CC PDB; 1LFI; 31-OCT-93.  
CC PDB; 1LGB; 31-AUG-94.  
CC PDB; 1LGC; 31-AUG-94.  
CC PDB; 1BKA; 08-NOV-96.  
CC PDB; 1DSN; 08-MAR-96.  
CC PDB; 1HSE; 12-MAR-97.  
CC PDB; 1VFD; 21-APR-97.

```
Query Match          94.4%; Score 67; DB 1; Length 711;
Best Local Similarity 91.7%; Pred. No. 0.0002;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWKRNRKVR 12
   |||||:|||||
DB 39 CFQWRNRKVR 50

RESULT 2
TRFL CAMDR
ID TRFL CAMDR STANDARD; PRT; 708 AA.
AC QPTUMQ; QWZS5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lactotransferrin precursor (Lactoferrin).
GN LTF.
OS Camelus dromedarius (Dromedary) (Arabic camel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.
OX NCBI_TaxID=9838;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sonali; TISSUE=Lactating mammary gland;
RA Kapteer S.R.; Ackermann M.; Farah Z.; Puhon Z.;
RT "Sequence analysis of camel (Camelus dromedarius) lactoferrin.";
RL Int. Dairy J. 9:481-486(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Paramasivam M.; Srinivasan A.; Singh R.; Sahani M.S.; Singh T.P.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE (BY SIMILARITY).
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AJ131674; CAB53387.1; -.
DR EMBL; AF165879; AAF82241.1; -.
DR HSSP; O77811; 1B1X.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 2.
DR PRINTS; PR00422; TRANSFERRIN.
DR SMART; SM00094; TR_FER; 2.
DR PROSITE; PS00205; TRANSFERRIN 1; 2.
DR PROSITE; PS00206; TRANSFERRIN 2; 2.
DR PROSITE; PS00207; TRANSFERRIN 3; 2.
KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;
Signal.
FT CHAIN 1 19 BY SIMILARITY.
FT SIGNAL 20 708 LACTOTRANSFERRIN.
FT REPEAT 20 363 1.
FT REPEAT 364 708 2.
FT DISULFID 28 64 BY SIMILARITY.
FT DISULFID 38 55 BY SIMILARITY.
FT DISULFID 134 217 BY SIMILARITY.
FT DISULFID 176 192 BY SIMILARITY.
FT DISULFID 189 200 BY SIMILARITY.
FT DISULFID 250 264 BY SIMILARITY.
FT DISULFID 367 399 BY SIMILARITY.
```

```
FT DISULFID 377 390 BY SIMILARITY.
FT DISULFID 424 703 BY SIMILARITY.
FT DISULFID 444 666 BY SIMILARITY.
FT DISULFID 476 551 BY SIMILARITY.
FT DISULFID 500 694 BY SIMILARITY.
FT DISULFID 510 524 BY SIMILARITY.
FT DISULFID 521 534 BY SIMILARITY.
FT DISULFID 592 606 BY SIMILARITY.
FT DISULFID 644 649 BY SIMILARITY.
FT METAL 79 79 IRON 1 (BY SIMILARITY).
FT METAL 111 111 IRON 1 (BY SIMILARITY).
FT METAL 211 211 IRON 1 (BY SIMILARITY).
FT METAL 272 272 IRON 1 (BY SIMILARITY).
FT METAL 414 414 IRON 2 (BY SIMILARITY).
FT METAL 452 452 IRON 2 (BY SIMILARITY).
FT METAL 545 545 IRON 2 (BY SIMILARITY).
FT METAL 614 614 IRON 2 (BY SIMILARITY).
FT BINDING 140 140 ANION (BY SIMILARITY).
FT BINDING 482 482 ANION (BY SIMILARITY).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 385 385 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 537 537 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 594 594 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 261 261 F -> S (IN REF. 2).
FT CONFLICT 304 304 G -> A (IN REF. 2).
FT CONFLICT 330 330 S -> P (IN REF. 2).
FT CONFLICT 492 494 LLS -> PLF (IN REF. 2).
FT CONFLICT 506 506 L -> F (IN REF. 2).
FT CONFLICT 609 609 A -> P (IN REF. 2).
FT CONFLICT 642 642 R -> Q (IN REF. 2).
SQ SEQUENCE 708 AA; 77211 MW; 0B0C175A0B69D430 CRC64;

Query Match          70.4%; Score 50; DB 1; Length 708;
Best Local Similarity 66.7%; Pred. No. 0.22;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWKRNRKVR 12
   |||||:|||||
DB 38 CAQWRNRKVR 49

RESULT 3
TRFL CAPHI
ID TRFL CAPHI STANDARD; PRT; 708 AA.
AC Q29477; Q29479;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lactotransferrin precursor (lactoferrin).
GN LTF.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RP TISSUE=Mammary gland;
RA Lee T.; Yu S.; Kim S.; Lee K.; Yu D.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RC MEDLINE=94380047; PubMed=8093048;
RA le Provost F.; Nocart M.; Guerin G.; Martin P.;
RT "Characterization of the goat lactoferrin cDNA. Assignment of the
RT relevant locus to bovine U12 syntenic group.";
RL Biochem. Biophys. Res. Commun. 203:1324-1332(1994).
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
```

CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U53857; AAA97958.1; --  
 CC EMBL; X78902; CAA55517.1; --  
 CC HSP; O77698; 1CE2.  
 CC InterPro; IPR001156; Transferrin.  
 CC Pfam; PF00405; transferrin; 2.  
 CC PRINTS; PR00422; TRANSFERRIN.  
 CC SMART; SM00094; TR FER; 2.  
 CC PROSITE; PS00205; TRANSFERRIN\_1; 2.  
 CC PROSITE; PS00206; TRANSFERRIN\_2; 2.  
 CC PROSITE; PS00207; TRANSFERRIN\_3; 2.  
 CC Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 CC Signal.  
 CC -----  
 CC SIGNAL 1 19 BY SIMILARITY.  
 CC CHAIN 20 708 LACTOTRANSFERRIN.  
 CC REPEAT 20 363 1.  
 CC REPEAT 364 708 2.  
 CC DISULFID 28 64 BY SIMILARITY.  
 CC DISULFID 38 55 BY SIMILARITY.  
 CC DISULFID 134 217 BY SIMILARITY.  
 CC DISULFID 176 192 BY SIMILARITY.  
 CC DISULFID 189 200 BY SIMILARITY.  
 CC DISULFID 250 264 BY SIMILARITY.  
 CC DISULFID 367 399 BY SIMILARITY.  
 CC DISULFID 377 390 BY SIMILARITY.  
 CC DISULFID 424 703 BY SIMILARITY.  
 CC DISULFID 444 666 BY SIMILARITY.  
 CC DISULFID 476 551 BY SIMILARITY.  
 CC DISULFID 500 694 BY SIMILARITY.  
 CC DISULFID 510 524 BY SIMILARITY.  
 CC DISULFID 521 534 BY SIMILARITY.  
 CC DISULFID 592 606 BY SIMILARITY.  
 CC DISULFID 644 649 BY SIMILARITY.  
 CC METAL 79 IRON 1 (BY SIMILARITY).  
 CC METAL 111 IRON 1 (BY SIMILARITY).  
 CC METAL 211 IRON 1 (BY SIMILARITY).  
 CC METAL 272 IRON 1 (BY SIMILARITY).  
 CC METAL 414 414 IRON 2 (BY SIMILARITY).  
 CC METAL 452 452 IRON 2 (BY SIMILARITY).  
 CC METAL 545 545 IRON 2 (BY SIMILARITY).  
 CC METAL 614 614 IRON 2 (BY SIMILARITY).  
 CC BINDING 140 140 ANION (BY SIMILARITY).  
 CC BINDING 482 482 ANION (BY SIMILARITY).  
 CC CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 495 495 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 564 564 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CONFLICT 56 56 I -> V (IN REF. 2).  
 CC CONFLICT 88 88 L -> R (IN REF. 2).  
 CC CONFLICT 124 124 Q -> K (IN REF. 2).  
 CC CONFLICT 154 154 F -> P (IN REF. 2).  
 CC CONFLICT 304 304 S -> R (IN REF. 2).  
 CC CONFLICT 414 414 D -> G (IN REF. 2).  
 CC SEQUENCE 708 AA; 77358 MW, P2EDA3C8339960D CRC64;  
 CC -----  
 CC Query Match 70.4%; Score 50; DB 1; Length 708;  
 CC Best Local Similarity 63.6%; Pred. No. 0.22;  
 CC Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 CC -----  
 CC 1 CFQWKNRKRKV 11  
 CC 38 CQWQRRKRKL 48

RESULT 4  
 TRFL HORSE  
 ID TRFL HORSE STANDARD; PRT; 695 AA.  
 AC 077811;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin) (Fragment).  
 GN LTF.  
 OS Equus caballus (Horse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCBI\_TaxID=9796;  
 RN (1)\_TaxID=9796;  
 RP SEQUENCE FROM N.A.  
 RA Paramasivam M., Srinivasan A., Yadav M.P., Singh T.P.;  
 RT "cDNA sequence of mare lactoferrin";  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).  
 RC TISSUE=Milk;  
 RX MEDLINE=9296631; PubMed=10366507;  
 RA Sharma A.K., Paramasivam M., Srinivasan A., Yadav M.P., Singh T.P.;  
 RT "Three-dimensional structure of mare diferric lactoferrin at 2.6-A  
 resolution.";  
 RL J. Mol. Biol. 289:303-317(1999).  
 CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY BICARBONATE.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AJ010930; CAA09407.1; --  
 CC FDB; 1BX; 02-DEC-98.  
 CC FDB; 1B7U; 02-FEB-99.  
 CC FDB; 1B7Z; 02-FEB-99.  
 CC InterPro; IPR001156; Transferrin.  
 CC Pfam; PF00405; transferrin; 2.  
 CC PRINTS; PR00422; TRANSFERRIN.  
 CC SMART; SM00094; TR FER; 2.  
 CC PROSITE; PS00205; TRANSFERRIN\_1; 2.  
 CC PROSITE; PS00206; TRANSFERRIN\_2; 2.  
 CC PROSITE; PS00207; TRANSFERRIN\_3; 1.  
 CC Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 CC Signal; 3D-structure.  
 CC NON\_TER 1 1  
 CC SIGNAL <1 6  
 CC CHAIN 7 695 LACTOTRANSFERRIN.  
 CC REPEAT 7 350 1.  
 CC REPEAT 351 695 2.  
 CC DISULFID 15 51  
 CC DISULFID 25 42  
 CC DISULFID 121 204  
 CC DISULFID 163 179  
 CC DISULFID 166 189  
 CC DISULFID 176 187  
 CC DISULFID 237 251  
 CC DISULFID 354 386  
 CC DISULFID 364 377  
 CC DISULFID 411 690  
 CC DISULFID 431 653

FT DISULFID 463 538  
 FT DISULFID 467 681  
 FT DISULFID 497 511  
 FT DISULFID 508 521  
 FT DISULFID 579 593  
 FT DISULFID 631 636  
 FT METAL 66 66  
 FT METAL 98 98  
 FT METAL 198 198  
 FT METAL 259 259  
 FT METAL 401 401  
 FT METAL 439 439  
 FT METAL 532 532  
 FT METAL 601 601  
 FT BINDING 127 127  
 FT BINDING 469 469  
 FT CARBOHYD 143 143  
 FT CARBOHYD 287 287  
 FT CARBOHYD 482 482  
 SQ SEQUENCE 695 AA; 75991 MW; 07BB84D50E1B165D CRC64;

Query Match 59.2%; Score 42; DB 1; Length 695;  
 Best Local Similarity 58.3%; Pred. No. 5.7;  
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQKRNKVR 12  
 Db 25 CAKFNKVR 36

## RESULT 5

RL28\_THETN STANDARD; PRT; 62 AA.  
 AC Q8F9U1;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE 50S ribosomal protein L28.  
 GN RPMB OR TPE1495.  
 OS Thermoaerobacter tengcongensis.  
 CC Bacteria; Firmicutes; Clostridia; Thermoaerobacteriales;  
 CC Thermoaerobacteriaceae; Thermoaerobacter.  
 OX NCBI\_TaxID=119072;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MB4 / JCM 11007;  
 RX MEDLINE=21992816; PubMed=11997336;  
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
 RA Tan H., Chen R., Wang J., Yu J., Yang H.,  
 RT "A complete sequence of T. tengcongensis genome."  
 RL Genome Res. 12:689-700(2002).  
 CC -1- SIMILARITY: BELONGS TO THE L28P FAMILY OF RIBOSOMAL PROTEINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).)  
 CC  
 DR EMBL; AF013107; AM24713.1; -  
 DR Ribosomal protein; Complete proteome.  
 KW SEQUENCE 62 AA; 7037 MW; CAADB605C81D495D CRC64;

Query Match 57.7%; Score 41; DB 1; Length 62;  
 Best Local Similarity 70.8%; Pred. No. 0.69;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 QWKNRKRVR 12  
 Db 27 RWKNRKRVR 36

## RESULT 6

TRFL\_MOUSE STANDARD; PRT; 707 AA.  
 AC P08071; P70690; Q61799; Q922P2;  
 DT 01-AUG-1998 (Rel. 08, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE Lactotransferrin precursor (Lactoferrin).  
 GN LTF.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Uterus;  
 RX MEDLINE=87280033; PubMed=3611056;  
 RA Pentecost B.T., Teng C.T.;  
 RT "Lactotransferrin is the major estrogen inducible protein of mouse  
 RT uterine secretions."  
 RL J. Biol. Chem. 262:10134-10139(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Uterus;  
 RA Morishita K.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 1-14 FROM N.A.  
 RX MEDLINE=92042099; PubMed=1939212;  
 RA Liu Y., Teng C.T.;  
 RT "Characterization of estrogen-responsive mouse lactoferrin promoter."  
 RL J. Biol. Chem. 266:21880-21885(1991)  
 CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY BICARBONATE.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).)  
 CC  
 DR EMBL; J03298; AAA40525.1; -  
 DR EMBL; D88510; BAA13633.1; -  
 DR EMBL; BC006904; AAH06904.1; -  
 DR EMBL; M74778; AAA39427.1; -  
 DR PIR; A28438; A28438.  
 DR HSP; P02788; ICB6.  
 DR MGD; MGI:96837; Ltf.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 2.  
 DR PRINTS; PRO0422; TRANSFERRIN.  
 DR SMART; SM00094; TR\_FER; 2.  
 DR PROSITE; PS00205; TRANSFERRIN\_1; 1.  
 DR PROSITE; PS00206; TRANSFERRIN\_2; 2.  
 DR PROSITE; PS00207; TRANSFERRIN\_3; 2.  
 KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 KW Signal.  
 FT SIGNAL 1 19 BY SIMILARITY.  
 FT CHAIN 20 707 LACTOTRANSFERRIN.  
 FT REPEAT 20 357 1.

FT REPEAT 358 707 2.  
 FT DISULFID 27 63 BY SIMILARITY.  
 FT DISULFID 37 54 BY SIMILARITY.  
 FT DISULFID 133 216 BY SIMILARITY.  
 FT DISULFID 175 191 BY SIMILARITY.  
 FT DISULFID 188 199 BY SIMILARITY.  
 FT DISULFID 249 263 BY SIMILARITY.  
 FT DISULFID 366 398 BY SIMILARITY.  
 FT DISULFID 376 389 BY SIMILARITY.  
 FT DISULFID 423 702 BY SIMILARITY.  
 FT DISULFID 443 665 BY SIMILARITY.  
 FT DISULFID 475 550 BY SIMILARITY.  
 FT DISULFID 499 693 BY SIMILARITY.  
 FT DISULFID 509 523 BY SIMILARITY.  
 FT DISULFID 520 533 BY SIMILARITY.  
 FT DISULFID 591 605 BY SIMILARITY.  
 FT DISULFID 643 648 BY SIMILARITY.  
 FT METAL 78 78 IRON 1 (BY SIMILARITY).  
 FT METAL 110 110 IRON 1 (BY SIMILARITY).  
 FT METAL 210 210 IRON 1 (BY SIMILARITY).  
 FT METAL 271 271 IRON 1 (BY SIMILARITY).  
 FT METAL 413 413 IRON 2 (BY SIMILARITY).  
 FT METAL 451 451 IRON 2 (BY SIMILARITY).  
 FT METAL 544 544 IRON 2 (BY SIMILARITY).  
 FT METAL 613 613 IRON 2 (BY SIMILARITY).  
 FT BINDING 139 139 ANION (POTENTIAL).  
 FT BINDING 481 481 ANION (POTENTIAL).  
 FT CARBOHYD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 494 494 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 1 2 MR -> IQG (IN REF. 1).  
 FT CONFLICT 25 25 R -> Q (IN REF. 2).  
 FT CONFLICT 82 82 M -> L (IN REF. 2).  
 FT CONFLICT 359 359 S -> T (IN REF. 2).  
 FT CONFLICT 382 382 A -> D (IN REF. 1).  
 FT CONFLICT 449 449 E -> G (IN REF. 2).  
 FT CONFLICT 629 629 L -> V (IN REF. 1).  
 FT CONFLICT 707 707 AA; 77865 MW; F26AE0340A4C19A8 CRC64;  
 SQ SEQUENCE

Query Match 57.7%; Score 41; DB 1; Length 707;  
 Best Local Similarity 54.5%; Pred. No. 8.7;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 CFQWKNRNVK 11  
 DB 37 CLRQWENRNV 47

RESULT 7  
 ID\_DYHB CHLRE STANDARD; PRT; 4568 AA.  
 AC Q3955;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Dynein beta chain, flagellar outer arm.  
 GN ODA4 OR ODA-4 OR SUP1.  
 OS Chlamydomonas reinhardtii.  
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
 OC Chlamydomonadaceae; Chlamydomonas.  
 OX NCBI\_TaxID=3055;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=21GR;  
 RX MEDLINE=94274778; PubMed=8006077;  
 RA Michell D.R.; Brown K.S.;  
 RT "Sequence analysis of the Chlamydomonas alpha and beta dynein heavy chain genes."  
 RL J. Cell Sci. 107:635-644(1994).  
 CC -1- FUNCTION: FORCE GENERATING PROTEIN OF EUKARYOTIC CILIA AND FLAGELLA. PRODUCES FORCE TOWARDS THE MINUS ENDS OF MICROTUBULES.  
 CC DYNEIN HAS ATPASE ACTIVITY.  
 CC -1- SUBUNIT: CONSISTS OF AT LEAST 3 HEAVY CHAINS (ALPHA, BETA AND GAMMA), 2 INTERMEDIATE CHAINS AND 8 LIGHT CHAINS.

CC -1- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; U02963; AAA19956.1; -  
 DR InterPro; IPR004273; Dynein\_heavy.  
 DR Pfam; PF03028; Dynein\_heavy; 1.  
 DR Motor protein; Microtubules; Dynein; ATP-binding; Flagella;  
 KW Coiled coil.  
 FT DOMAIN 277 293 COILED COIL (POTENTIAL).  
 FT DOMAIN 1158 1175 COILED COIL (POTENTIAL).  
 FT DOMAIN 1372 1400 COILED COIL (POTENTIAL).  
 FT DOMAIN 1614 1650 COILED COIL (POTENTIAL).  
 FT DOMAIN 1778 1825 COILED COIL (POTENTIAL).  
 FT DOMAIN 2017 2045 MICROTUBULE-BINDING (POTENTIAL).  
 FT DOMAIN 2831 2848 COILED COIL (POTENTIAL).  
 FT DOMAIN 3106 3162 COILED COIL (POTENTIAL).  
 FT DOMAIN 3339 3425 COILED COIL (POTENTIAL).  
 FT DOMAIN 3648 3728 COILED COIL (POTENTIAL).  
 FT NP\_BIND 1919 1926 ATP (POTENTIAL).  
 FT NP\_BIND 2202 2209 ATP (POTENTIAL).  
 FT NP\_BIND 2530 2537 ATP (POTENTIAL).  
 FT NP\_BIND 2879 2886 ATP (POTENTIAL).  
 SQ SEQUENCE 4568 AA; 519961 MW; 9A9A5393C7C36AE7 CRC64;

Query Match 56.3%; Score 40; DB 1; Length 4568;  
 Best Local Similarity 41.7%; Pred. No. 92;  
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 CFQWKNRNVK 12  
 DB 1852 CFQWQSLRYIQ 1863

RESULT 8  
 ID\_NLA DROME STANDARD; PRT; 292 AA.  
 AC Q9XZL8; Q9V391;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Nebula protein.  
 GN NLA OR CG6072.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA McCormick A.V.; Goldberg M.L.;  
 RT "Gene required for elongation of meiosis I spindle in Drosophila females."  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkely;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D.; Celniker S.E.; Holt R.A.; Evans C.A.; Gocayne J.D.;  
 RA Anantides P.G.; Scherer S.E.; Li P.W.; Hoskins R.A.; Galie R.F.;  
 RA Sutton G.G.; Wortman J.R.; Randell M.D.; Zhang Q.; Chen L.X.;  
 RA Brandon R.C.; Rogers Y.-H.C.; Blazej R.G.; Champe M.; Pfeiffer B.D.;  
 RA Wan K.H.; Doyle C.; Baxter E.G.; Helt G.; Nelson C.R.; Miklos G.L.G.;  
 RA Abril J.F.; Agbayani A.; An H.-J.; Andrews-Frankoch C.; Baldwin D.;  
 RA Balow R.M.; Basu A.; Baxendale J.; Bayraktaroglu L.; Beasley E.M.;  
 RA Beeson K.Y.; Benos P.V.; Berman B.P.; Bhandari D.; Bolshakov S.;

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Farrar C., Ferriera S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.N., Glasser K.,  
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M.E., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laeko P., Lei Y., Levitsky A.P., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.F., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nussekern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskae R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RA "The genome sequence of *Drosophila melanogaster*."  
 RT Science 287:2185-2195(2000).  
 RL  
 CC -!- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF147700; AAD33987.1; -;  
 DR EMBL; AE003712; AAF5285.1; -;  
 DR Flybase; FBgn0026629; nla.  
 SQ SEQUENCE 292 AA; 31423 MW; 64F1BBF5F6EA6CF9 CRC64;  
 Query Match 54.9%; Score 39; DB 1; Length 292;  
 Best Local Similarity 54.5%; Pred. No. 7.9; Indels 0; Gaps 0;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 FOWLRMRKVR 12  
 DB 150 FOWLRFRRLR 160  
 RESULT 9  
 ID RL28\_BACST STANDARD; PRT; 60 AA.  
 AC P23374;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 508 ribosomal protein L28.  
 GN RPB.  
 OS Bacillus stearothermophilus.  
 OC Bacteria; Firmicutes; Bacillales; Geobacillus.  
 CC NCBI TaxID=1422;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92075759; PubMed=1742360;  
 RA Krufft V., Kapp U., Wittmann-Liebold B.;  
 RT "Characterization and primary structure of proteins L28, L33 and L34  
 RT from *Bacillus stearothermophilus* ribosomes.";  
 RL Biochimie 73:855-860(1991).

CC -!- SIMILARITY: BELONGS TO THE L28P FAMILY OF RIBOSOMAL PROTEINS.  
 DR PIR; A48396; A48396.  
 DR InterPro; IPR001383; Ribosomal L28.  
 DR Pfam; PF00830; Ribosomal L28; 1.  
 DR TIGRFAMs; TIGR00009; L28; 1.  
 KW Ribosomal protein.  
 FT INIT MET 0  
 SQ SEQUENCE 60 AA; 6810 MW; 2AD9161CD60B82F4 CRC64;  
 Query Match 53.5%; Score 38; DB 1; Length 60;  
 Best Local Similarity 66.7%; Pred. No. 2.3;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 WGRNMRKVR 12  
 DB 27 WGRNLRKVR 35  
 RESULT 10  
 ID RL28\_LISMO STANDARD; PRT; 62 AA.  
 AC Q92AJ2;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 508 ribosomal protein L28.  
 GN RPB OR LMO1816 OR LMO1930.  
 OS Listeria monocytogenes, and  
 OS Listeria innocua.  
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 CC NCBI TaxID=1639, 1642;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=L.monocytogenes, and L.innocua;  
 RC STRAIN=EG-e / Serovar 1/2a, and CLIP 11262 / Serovar 6a;  
 RX MEDLINE=21537279; PubMed=11679669;  
 RA Glaser P., Francuel P., Blocher H., Brandt P., Chakraborty T.,  
 RA Baquero F., Berche P., Blocher H., de Daruvar A., Dehoux P.,  
 RA Charbit A., Cherouani F., Couve E., de Daruvar A., Dussurget O.,  
 RA Donann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
 RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,  
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
 RA Jones L.-M., Kaerst J., Kref J., Kuhn M., Kunst F., Kurapkat G.,  
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,  
 RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,  
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tisseret A.,  
 RA Vaquer-Solana J.-A., Voss H., Wehland J., Cossart P.;  
 RT "Comparative genomics of *Listeria* species.";  
 RL Science 294:849-852(2001).  
 CC -!- SIMILARITY: BELONGS TO THE L28P FAMILY OF RIBOSOMAL PROTEINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AL591981; CAC99894.1; -;  
 DR EMBL; AL596170; CAC97160.1; -;  
 DR Listlist; LMO1930; -;  
 DR Listlist; LMO01816; -;  
 DR InterPro; IPR001383; Ribosomal L28.  
 DR Pfam; PF00830; Ribosomal L28; 1.  
 DR TIGRFAMs; TIGR00009; L28; 1.  
 KW Ribosomal protein; Complete proteome.  
 RP SEQUENCE 62 AA; 6991 MW; AA43DE039213C562 CRC64;  
 Query Match 53.5%; Score 38; DB 1; Length 62;  
 Best Local Similarity 66.7%; Pred. No. 2.3;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

QY 4 WKRNKVR 12
DB 29 WKANLQVR 37

RESULT 11
RPOB LIBAF
ID RPOB LIBAF STANDARD; PRT; 146 AA.
AC P41187;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA-directed RNA polymerase beta chain (BC 2.7.7.6) (Transcriptase
  beta chain) (RNA polymerase beta subunit) (Fragment).
GN RPOB.
OS Liberibacter africanus (Liberibacter africanum).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Candidatus Liberibacter.
OX NCBI_TaxID=34020;
RN [1]
RC SEQUENCE FROM N.A.
RA Planet P., Jagoueix S., Bove J.M., Garnier M.;
RT "Detection and characterization of the African Citrus Greening
  Liberibacter by amplification, cloning and sequencing of the rplKAL-
  rpoBC operon.";
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
  OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
  SUBSTRATES.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
  (RNA)(N).
CC -!- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
  ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
  BETA' CHAIN.
CC -!- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
CC EMBL; U03675; AAA19557.1; -
CC InterPro; IPR001572; RNA_pol_B.
CC Pfam; PF00562; RNA_pol_B; 1.
CC PROSITE; PS01166; RNA_POL_BETA; PARTIAL.
CC Transferase; Transcription; DNA-directed RNA polymerase.
FT NON TER 146 146
SQ SEQUENCE 146 AA; 16731 MW; 815D6EDD68FD8943 CRC64;

Query Match 53.5%; Score 38; DB 1; Length 146;
Best Local Similarity 60.0%; Pred. No. 5.7;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWKNRKR 10
DB 10 CVQWSRGARK 19

RESULT 12
Y125 MYCCA
ID Y125 MYCCA STANDARD; PRT; 267 AA.
AC P33661;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical protein (Fragment).
OS Mycoplasma capricolum
OC Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;
OC Entomoplasmataceae.

```

```

OX NCBI_TaxID=2095;
RN [1]
RC SEQUENCE FROM N.A.
RA STRAIN=ATCC 27343 / KID;
RX MEDLINE=96059641; PubMed=7476192;
RA Bork P., Ouzounis C., Casari G., Schneider R., Sander C.,
  Dolan M., Gilbert W., Gillevet P.M.;
RT "Exploring the Mycoplasma capricolum genome: a minimal cell reveals
  its physiology.";
RL Mol. Microbiol. 16:955-967(1995).
CC -!- SIMILARITY: BELONGS TO THE COF/YEHA/YIDA/YIGL (E. COLL) / YCSE/YXEH
  (B. SUBTILIS) FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
CC EMBL; Z33006; CAA83689.1; -
CC InterPro; IPR001454; Hlgase/hydrilase.
CC InterPro; IPR000150; Hypothet_cof.
CC Pfam; PF0702; Hydrolase; 1.
CC PROSITE; PS01228; COF 1; 1.
CC PROSITE; PS01229; COF 2; 1.
CC Hypothetical protein.
FT NON TER 267 267
SQ SEQUENCE 267 AA; 30425 MW; D5912DD5B39A8451 CRC64;

Query Match 53.5%; Score 38; DB 1; Length 267;
Best Local Similarity 58.3%; Pred. No. 11;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWKNRKR 12
DB 158 CFQWKNRQMR 169

RESULT 13
VIF SIVS4
ID VIF SIVS4 STANDARD; PRT; 214 AA.
AC P12505;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DE Viron infectivity factor (SOR protein) (Q protein).
GN VIF.
OS Simian immunodeficiency virus (F236/smH4 isolate) (sooty mangabey).
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11737;
RN [1]
RC SEQUENCE FROM N.A.
RX MEDLINE=89262053; PubMed=2786147;
RA Hirsch V.M., Olmstead R.A., Murphy-Corb M., Purcell R.H.,
  Johnson P.R.;
RT "An African primate lentivirus (SIVem) closely related to HIV-2.";
RL Nature 339:389-392(1989). VIRUS INFECTIVITY.
CC -!- FUNCTION: DETERMINES VIRUS INFECTIVITY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
CC EMBL; X14307; CAA32484.1; -
CC PIR; S07989; S07989.
CC HIV; X14307; VIF$SMH4.
CC InterPro; IPR000475; Viral_infect.

```



```

DR Pfam; PF00559; Vif; 1.
DR PRINTS; PR00349; VIRIONINFECT.
DR ProDom; PD000063; Viral_infect; 1.
SQ AIDS.
SQ SEQUENCE 214 AA; 25140 MW; 9BCE3884EC454BF3D CRC64;

Query Match 52.1%; Score 37; DB 1; Length 214;
Best Local Similarity 75.0%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 0; Gaps 0;

Qy 3 QWRNRK 10
|:|:|:|:|
Db 170 QWRNRK 177

RESULT 14
YNR2_CAEEL STANDARD; PRT; 783 AA.
AC Q21986;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein R13G10.2 in chromosome III.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gardner A.E.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Durbin R.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- COFACTOR: FAD (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE FLAVIN MONOAMINE OXIDASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; Z35602; CA84671.2; -.
DR WormPep; R13G10.2; CE25088.
DR InterPro; IPR02937; Amino_oxidase.
DR Pfam; PF01593; Amino_oxidase; 1.
KW Hypothetical protein; Oxidoreductase; Flavoprotein; FAD.
FT NP_BIND 311 365 FAD (ADP PART) (POTENTIAL).
SQ SEQUENCE 783 AA; 88799 MW; 8D087E96464DC908 CRC64;

Query Match 52.1%; Score 37; DB 1; Length 783;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CFQWKNRKVR 12
|:|:|:|:|
Db 540 CIDWGRDRKVK 551

RESULT 15
1A34_HUMAN STANDARD; PRT; 365 AA.
AC P30453; P30454;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE HLA class I histocompatibility antigen, AW-34(A-10) alpha chain

```

```

DE GN HLA-A OR HLA-A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (A*3401/A*3402).
RX MEDLINE=93056508; PubMed=143115;
RA Madrigal J.A., Belich M.P., Hildebrand W.H., Benjamin R.J.,
RA Little A.-M., Zemmour J., Ennis P.D., Ward F.E., Petzl-Erler M.L.,
RA Martell R.W., du Toit E.D., Parham P.;
RT "Distinctive HLA-A,B antigens of black populations formed by
RT interallelic conversion."
RL J. Immunol. 149:3411-3415(1992).
RN [2]
RP SEQUENCE FROM N.A. (A*3401/A*3402).
RX MEDLINE=9323211; PubMed=8475492;
RA Madrigal J.A., Hildebrand W.H., Belich M.P., Benjamin R.J.,
RA Little A.-M., Zemmour J., Ennis P.D., Ward F.E., Petzl-Erler M.L.,
RA du Toit E.D., Parham P.;
RT "Structural diversity in the HLA-A10 family of alleles: correlations
RT with serology."
RL Tissue Antigens 41:72-80(1993).
CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -1- POLYMORPHISM: THE FOLLOWING ALLELES OF AW-34 ARE KNOWN: A*3401
CC (AW-34.1) AND A*3402 (AW34.2). THE SEQUENCE SHOWN IS THAT OF
CC A*3401.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X61704; CAA43873.1; -.
DR EMBL; X61705; CAA43874.1; -.
DR PIR; S16767; S16767.
DR PIR; S16771; S16771.
DR HSP; O19673; 1HSB.
DR MIM; 142800; -.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR ProDom; PD000050; MHC_I; 1.
DR SMART; SM00407; IGL1; 1.
DR PROSITE; PS00250; IG_MHC; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal; Polymorphism.
FT CHAIN 1 24
FT SIGNAL 25 365 HLA CLASS I HISTOCOMPATIBILITY ANTIGEN,
FT AW-34 (A-10) ALPHA CHAIN.
FT DOMAIN 25 114 EXTRACELLULAR ALPHA-1.
FT DOMAIN 115 206 EXTRACELLULAR ALPHA-2.
FT DOMAIN 207 298 EXTRACELLULAR ALPHA-3.
FT DOMAIN 299 308 CONNECTING PEPTIDE.
FT TRANSMEM 309 332
FT DOMAIN 333 365
FT CARBOHYD 110 110
FT DISULFID 125 188
FT DISULFID 227 283
FT VARIANT 3 3 I -> V (IN A*3402).
FT VARIANT 90 90 K -> N (IN A*3402).
FT VARIANT 121 121 R -> I (IN A*3402).
FT VARIANT 121 121 /FTID=VAR_004381.
FT /FTID=VAR_004381.

```

```

FT VARIANT 129 129 P -> S (IN A*3402).
FT FTID=VAR 004382.
FT VARIANT 138 138 Q -> R (IN A*3402).
FT FTID=VAR 004383.
FT VARIANT 180 180 W -> L (IN A*3402).
FT FTID=VAR 004384.
FT VARIANT 312 312 L -> I (IN A*3402).
FT FTID=VAR 004385.
SQ SEQUENCE 365 AA; 41055 MW; 063BF63E636E01F6 CRC64;

Query Match 50.7%; Score 36; DB 1; Length 365;
Best Local Similarity 66.7%; Pred. NO. 34;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 WKRNRKVR 12
Db 84 WDRNRKVK 92

```

Search completed: February 21, 2003, 07:51:39  
 Job time : 6.2093 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

3M protein - protein search, using sw model

Run on: February 21, 2003, 07:44:43 ; Search time 21.6744 Seconds  
(without alignments)  
114.078 Million cell updates/sec

Title: US-09-743-107B-92

Perfect score: 71

Sequence: 1 CFQWKNRKYR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 21.\*

1: sp archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mhc.\*

8: sp\_organelle.\*

9: sp\_phage.\*

10: sp\_plant.\*

11: sp\_rodent.\*

12: sp\_virus.\*

13: sp\_vertebrate.\*

14: sp\_unclassified.\*

15: sp\_rvirus.\*

16: sp\_bacteriap.\*

17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

result No.	Score	Match	Length	DB ID	Description
1	62	87.3	711	4	Q8TCD2 homo sapien
2	58	81.7	38	4	Q9UCY5 homo sapien
3	47	66.2	33	6	Q9TR80 ovis aries
4	45	63.4	275	5	Q93780 caenorhabdi
5	41	57.7	62	16	Q8R9U1 Q8R9U1 thermoanaer
6	41	57.7	205	16	Q986A0 rhizobium l
7	41	57.7	206	16	Q8UHC2 agrobacteri
8	41	57.7	206	16	Q92RH8 rhizobium m
9	41	57.7	208	16	Q8YFK3 bruceella me
10	41	57.7	511	16	Q8Z462 salmonella
11	40	56.3	81	15	Q90863 human immun
12	40	56.3	148	10	Q9XHP1 Q9XHP1 sesamum ind
13	40	56.3	274	4	Q96M21 Q96M21 homo sapien
14	40	56.3	584	17	Q9HPA3 halobacteri
15	39	54.9	273	2	O31090 rhizobium l
16	39	54.9	279	16	Q8XSE2 ralstonia s

```

17 39 54.9 447 5 017549
18 39 54.9 550 11 Q9JZS
19 39 54.9 759 16 Q9Z955
20 39 54.9 864 5 Q62582
21 39 54.9 864 5 Q8SRG3
22 39 54.9 864 5 Q8SQI6
23 38 53.5 81 15 Q90884
24 38 53.5 91 15 Q77855
25 38 53.5 105 10 Q9XFD5
26 38 53.5 207 10 Q9SML1
27 38 53.5 240 10 Q9SML2
28 38 53.5 253 12 Q68541
29 38 53.5 238 16 Q8YP77
30 38 53.5 306 4 Q8TAX2
31 38 53.5 329 2 Q9F7Y4
32 38 53.5 341 11 Q8R2A4
33 38 53.5 372 10 Q81653
34 38 53.5 393 10 Q9ZTP0
35 38 53.5 393 16 Q25611
36 38 53.5 402 10 Q9ZRH8
37 38 53.5 466 4 Q9NUS2
38 38 53.5 489 16 P96223
39 38 53.5 499 10 Q9XFX1
40 38 53.5 560 10 Q9FTT0
41 38 53.5 570 10 Q8S487
42 38 53.5 583 10 Q81437
43 38 53.5 662 12 Q9QU30
44 38 53.5 666 5 Q9W1D5
45 38 53.5 866 10 Q9FHI9

```

#### ALIGNMENTS

#### RESULT 1

```

Q8TCD2 PRELIMINARY; PRT; 711 AA.
ID AC Q8TCD2;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Lactotransferrin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PROSTATE;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC02347; AAH2347.1;
SQ SEQUENCE 711 AA; 78327 MW; 1B9C7BE097C45FAF CRC64;

```

Query Match 87.3%; Score 62; DB 4; Length 711;  
Best Local Similarity 90.9%; Pred. No. 0.0074;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFQWKNRKYR 11

Db 39 CFQWKNRKYR 49

#### RESULT 2

```

Q9UCY5 PRELIMINARY; PRT; 38 AA.
ID AC Q9UCY5;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Lactoferrin homolog (Fragment).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96081613; PubMed=8551695;  
 RA Sato I.;  
 RT "Characterization of the 84-kDa protein with ABH activity in human  
 serumal plasma.";  
 RL Jpn. J. Legal Med. 49:281-293 (1995).  
 DR HSP; P02788; IBA.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 1.  
 SQ SEQUENCE 38 AA; 4459 MW; 0402F49055EBDDDB CRC64;  
 Query Match 81.7%; Score 58; DB 4; Length 38;  
 Best Local Similarity 90.9%; Pred. No. 0.0019;  
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 FQWKNRMRKV 12  
 DB 21 FQWKNRMRKV 31  
 RESULT 3  
 Q9TR80  
 ID Q9TR80 PRELIMINARY; PRT; 33 AA.  
 AC Q9TR80;  
 DT 01-WAY-2000 (TREMBLrel. 13, Created)  
 DT 01-WAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE Lactoferrin (Fragment).  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=95127729; PubMed=7827104;  
 RA Qian Z.Y., Joiles P., Migliore-Sancur D., Fiat A.M.;  
 RL Biochim. Biophys. Acta 1243:25-32 (1995).  
 DR HSP; O77698; ICE2.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 1.  
 SQ SEQUENCE 33 AA; 3914 MW; D1904CAE15A73961 CRC64;  
 Query Match 66.2%; Score 47; DB 6; Length 33;  
 Best Local Similarity 54.5%; Pred. No. 0.16;  
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQWKNRMRKV 11  
 DB 19 CFQWKNRMRKV 29  
 RESULT 4  
 Q93780  
 ID Q93780 PRELIMINARY; PRT; 275 AA.  
 AC Q93780;  
 DT 01-FEB-1997 (TREMBLrel. 02, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE F53H4.4 protein.  
 GN F53H4.4  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX Dobson R.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=95069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology.";  
 RL Science 282:2012-2018 (1998).  
 DR EMBL; Z81089; CAB03137.1; -;  
 SQ SEQUENCE 275 AA; 31716 MW; 9D69524FFC704DB2 CRC64;  
 Query Match 63.4%; Score 45; DB 5; Length 275;  
 Best Local Similarity 72.7%; Pred. No. 3.1;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 FQWKNRMRKV 12  
 DB 262 FQWKNRMRKV 272  
 RESULT 5  
 Q8R9U1  
 ID Q8R9U1 PRELIMINARY; PRT; 62 AA.  
 AC Q8R9U1;  
 DT 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Ribosomal protein L28.  
 GN RPB OR TWE1495.  
 OS Thermoanaerobacter tengcongensis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
 OC Thermoanaerobacteriales; Thermoanaerobacteriaceae; Thermoanaerobacter.  
 OX NCBI\_TaxID=119072;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MB4T / JCM11007;  
 RX MEDLINE=21992816; PubMed=1197336;  
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;  
 RA "A complete sequence of T. tengcongensis genome.";  
 RT Genome Res. 12:689-700 (2002).  
 RL EMBL; AE013107; AAM24713.1; -;  
 KW Complete proteome.  
 SQ SEQUENCE 62 AA; 7037 MW; CAADB605C81D495D CRC64;  
 Query Match 57.7%; Score 41; DB 16; Length 62;  
 Best Local Similarity 70.0%; Pred. No. 3.6;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 QWKNRMRKV 12  
 DB 27 QWKNRMRKV 36  
 RESULT 6  
 Q986A0  
 ID Q986A0 PRELIMINARY; PRT; 205 AA.  
 AC Q986A0;  
 DT 01-OCT-2001 (TREMBLrel. 18, Created)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Pyridoxamine 5'-phosphate oxidase.  
 GN ML7454.  
 OS Rhizobium loti (Mesorhizobium loti).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Phyllobacteriaceae; Mesorhizobium.  
 OX NCBI\_TaxID=391;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MAFF303099;  
 RX Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
 RA Watanabe A., Iida K., Ishikawa K., Kawahara K., Kimura T.,  
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,

RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,  
 RA Takeuchi C., Yamada M., Tabata S.;  
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
 RT Mesorhizobium loti";  
 RL DNA Res. 7:331-338(2000).  
 DR EMBL; AP003011; BAB53553.1; -;  
 DR InterPro; IPR000659; Pyridox\_oxidase.  
 DR Pfam; PF01243; Pyridox\_oxidase; 1.  
 DR ProDom; PD008312; Pyridox\_oxidase; 1.  
 DR TIGRfam; TIGR00558; pdxH; 1.  
 DR PROSITE; PS01064; PYRIDOX\_OXIDASE; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 205 AA; 23300 MW; 0BAD54CD312327EA CRC64;

Query Match 57.7%; Score 41; DB 16; Length 205;  
 Best Local Similarity 58.3%; Pred. No. 12;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWKNRNRKVR 12  
 |||||  
 DB 88 CFHWKSLRRQVR 99

RESULT 7  
 Q8UHC2 PRELIMINARY; PRT; 206 AA.  
 IT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 YT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 YT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Pyridoxamine 5'-phosphate oxidase.  
 PE PDHX OR ATU0760 OR AGR C.1381.

NS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
 XC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 XC Rhizobiaceae; Rhizobium.  
 XX NCBI\_TaxID=176299;  
 XN [1]

RP SEQUENCE FROM N.A.  
 XX MEDLINE=21608550; PubMed=11743193;  
 YA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,  
 YA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,  
 YA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr., Woo L.,  
 YA Chapman P., Clendenning J., Decherage G., Gillet W., Grant C.,  
 YA Kutayav T., Levy R., Li M.-J., McClelland E., Palmeri A.,  
 YA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,  
 YA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
 YA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
 YA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
 YA Nester E.W.;  
 T "The genome of the natural genetic engineer Agrobacterium tumefaciens  
 T C58";  
 TL Science 294:2317-2323(2001).  
 [2]

FP SEQUENCE FROM N.A.  
 MEDLINE=21608551; PubMed=11743194;  
 XA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
 XA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,  
 XA Houlmiel K., Gordon J., Vaudin M., Tarchouk O., Epp A., Liu F.,  
 XA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,  
 XA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,  
 XA Cielo C., Slater S.;  
 T "Genome sequence of the plant pathogen and biotechnology agent  
 T Agrobacterium tumefaciens C58";  
 TL Science 294:2323-2328(2001).  
 R EMBL; AEO09043; AAL41776.1; -;  
 R EMBL; AEO08009; AAK86569.1; -;  
 W Complete proteome.  
 Q SEQUENCE 206 AA; 23720 MW; 3BE488AE5307C0C1 CRC64;

Query Match 57.7%; Score 41; DB 16; Length 206;  
 Best Local Similarity 58.3%; Pred. No. 12;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWKNRNRKVR 12  
 |||||  
 DB 88 CFHWKSLRRQVR 99

RESULT 8  
 Q92RH8 PRELIMINARY; PRT; 206 AA.  
 ID Q92RH8;  
 AC Q92RH8;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Probable pyridoxamine 5'-phosphate oxidase (PNP/OMP oxidase) protein  
 DE (EC 1.4.3.5).  
 DE PDHX OR R00895 OR SMC00069.  
 GN Rhizobium meliloti (Sinorhizobium meliloti).  
 OS Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Sinorhizobium.  
 CX NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021;  
 RX MEDLINE=21396507; PubMed=11481430;  
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,  
 RA Pohl T., Portetelle D., Fuehrer A., Fumelle B., Ramsperger U.,  
 RA Renard C., Thebault P., Vandenbol M., Weidner S., Gallibert F.;  
 RT "Analysis of the chromosome sequence of the legume symbiont  
 RT Sinorhizobium meliloti strain 1021";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 DR EMBL; AL591785; CAC45467.1; -;  
 DR InterPro; IPR000659; Pyridox\_oxidase.  
 DR Pfam; PF01243; Pyridox\_oxidase; 1.  
 DR ProDom; PD006312; Pyridox\_oxidase; 1.  
 DR TIGRfam; TIGR00558; pdxH; 1.  
 DR PROSITE; PS01064; PYRIDOX\_OXIDASE; UNKNOWN\_1.  
 DR Oxidoreductase; Complete proteome.  
 KW SEQUENCE 206 AA; 23900 MW; A2DB74229DACA97A CRC64;

Query Match 57.7%; Score 41; DB 16; Length 206;  
 Best Local Similarity 58.3%; Pred. No. 12;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWKNRNRKVR 12  
 |||||  
 DB 88 CFHWKSLRRQVR 99

RESULT 9  
 Q8YFK3 PRELIMINARY; PRT; 208 AA.  
 ID Q8YFK3;  
 AC Q8YFK3;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Probable pyridoxamine 5'-phosphate oxidase (EC 1.4.3.5).  
 GN BMEI1517.  
 OS Brucella melitensis.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Brucellaceae; Brucella.  
 CX NCBI\_TaxID=294559;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=16M / ATCC 23456 / BIOTYPE 1;  
 RX MEDLINE=20020109; PubMed=11756688;  
 RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Muier C., Los T.,  
 RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,  
 RA Jablonksi L., Larsen L., D'Souza M., Bernal A., Mazur M., Goltzman E.,  
 RA Salkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,  
 RA Haselkorn R., Kyrides N., Overbeek R.;  
 RT "The genome sequence of the facultative intracellular pathogen  
 RT Brucella melitensis";

RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448 (2002).

DR EMBL; AE009587; ALU52698.1; --  
DR InterPro; IPR000659; Pyridox oxidase.  
DR Pfam; PF01243; Pyridox oxidase; 1.  
DR ProDom; PD006312; Pyridox oxidase; 1.  
DR TIGRFAMs; TIGR00558; pdxh; 1.  
DR PROSITE; PS01064; PYRIDOX OXIDASE; 1.  
KW Oxidoreductase; Complete proteome; 1.  
SQ SEQUENCE 208 AA; 23866 MW; CBIF50BC9612DE28 CRC64;

Query Match 57.7%; Score 41; DB 16; Length 208;  
Best Local Similarity 58.3%; Pred. No. 12;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CFQWKNRNRKV 12

Db 90 CFHWKSLRRQVR 101

RESULT 10

Q8Z462 PRELIMINARY; PRT; 511 AA.  
AC Q8Z462;

DT 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Hypothetical protein STY3070.

GN STY3070

OS Salmonella typhi.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Salmonella.

OC NCBI\_TaxID=601;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CT18.

RX MEDLINE=21534947; PubMed=11677608;

RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,

RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,

RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,

RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,

RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,

RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,

RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,

RA Whitehead S., Barrrell B.G.;

RT "Complete genome sequence of a multiple drug resistant Salmonella

RL Nature 413:848-852 (2001)."

RL EMBL; AL627276; CAD06049.1; --

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 511 AA; 58126 MW; E2DDDL24E10D178B CRC64;

Query Match 57.7%; Score 41; DB 16; Length 511;  
Best Local Similarity 58.3%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CFQWKNRNRKV 12

Db 350 CFQWDMNKAQVR 361

RESULT 11

Q90863

ID Q90863 PRELIMINARY; PRT; 81 AA.

AC Q90863;

DT 01-NOV-1998 (TrEMBLrel. 08, Created)

DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Glycoprotein gp120 (Fragment).

GN ENV.

OS Human immunodeficiency virus type 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OC NCBI\_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NJS182;

RA Brandful J.A.M., Ampofo W.K., Janssens W., Adu-Sarkodie Y.,

RA Agyeyei F., Anyomi F., Aidoo S., Barnor J.S., Yamamoto N.,

RA Ishikawa K., Sata T., Kurata T.;

RT "Genetic and phylogenetic analysis of HIV-1 strains from Southern

RT Ghana".

RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ225659; CAA12541.1; --

DR InterPro; IPR000777; GPI20.

DR Pfam; PF00516; GPI20; 1.

KW AIDS; Coat protein; Glycoprotein.

FT NON\_TER 1

FT NON\_TER 81

SQ SEQUENCE 81 AA; 9138 MW; 2D43DCD554295572 CRC64;

Query Match 56.3%; Score 40; DB 15; Length 81;  
Best Local Similarity 56.7%; Pred. No. 7.1;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 QWKNRNV 11

Db 69 EWKELRV 77

RESULT 12

Q9XHP1

ID Q9XHP1 PRELIMINARY; PRT; 148 AA.

AC Q9XHP1;

DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE 2S albumin.

OS Sesamum indicum (Oriental sesame) (Gingelly).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; euasterids I; Lamiales; Pedaliaceae; Sesamum.

OC NCBI\_TaxID=4182;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=TAINAN 1;

RX MEDLINE=20074970; PubMed=10605554;

RA Tai S.S.K., Wu L.S.H., Chen E.C.F., Tzen J.T.C.;

RA "Molecular cloning of 11S globulin and 2S albumin, the two major seed

RT storage proteins in sesame."

RL J. Agric. Food Chem. 47:4932-4938 (1999).

DR EMBL; AF091841; AAD42943.1; --

DR InterPro; IPR003612; AAI.

DR InterPro; IPR000617; Napin.

DR InterPro; IPR001768; Try/amy1\_inhbr.

DR Pfam; PF00234; tryp\_alpha\_amy1; 1.

DR PRINTS; PR00496; NAPIN.

DR SMART; SM00499; AAI; 1.

SQ SEQUENCE 148 AA; 17524 MW; BA46B033BA13E3DE CRC64;

Query Match 56.3%; Score 40; DB 10; Length 148;  
Best Local Similarity 66.7%; Pred. No. 13;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQWKNRNR 9

Db 54 CMQWMSMR 62

RESULT 13

Q96M21

ID Q96M21 PRELIMINARY; PRT; 274 AA.

AC Q96M21;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE CDNA FLJ32891 fis, clone TEST12004929.

OS Homo sapiens (Human).

DC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
DD Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
DX NCBI\_TaxID=9606;  
DY [1]  
EZ SEQUENCE FROM N.A.  
FC TISSUE=TESTIS;  
GA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,  
GB Arita M., Mueashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,  
GC Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,  
GD Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,  
GE Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,  
GF Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,  
GH Sugano S., Nagahari K., Masuko Y., Nagai K., Isogai T.,  
GI "NEDO human cDNA sequencing project.";  
GT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AK057453; BAB71493.1; -.  
RQ SEQUENCE 274 AA; 30083 MW; 1DD3654D4135B2F CRC64;

Query Match 56.3%; Score 40; DB 4; Length 274;  
Best Local Similarity 50.0%; Pred. No. 25;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Y 1 CFQWKRNMKVR 12  
| | | | | | | | | |  
b 66 CFQWGRGVYLR 77

RESULT 14  
J9HPA3 PRELIMINARY; PRT; 584 AA.  
AC Q9HPA3;  
AD Q9HPA3;  
DT 01-MAR-2001 (TREMblrel. 16, Created)  
JT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
JT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
ZE VNGI732C.  
ZN VNGI732C.  
DS Halobacterium sp. (strain NRC-1).  
DC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
CC Halobacteriaceae; Halobacterium.  
DX NCBI\_TaxID=64091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20504483; PubMed=11016950;  
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,  
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,  
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,  
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Argevine C.M., Dale H.,  
RA Isenbarger T.A., Peck R.F., Fohlischroder M., Spudich J.L., Jung K.-H.,  
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Oner A.D.,  
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;  
RT "Genome sequence of Halobacterium species NRC-1.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
DR EMBL; AE005078; AAG19967.1; -.  
DR InterPro; IPR001646; Septide\_repeat.  
DR InterPro; IPR001622; K-channel\_pore.  
DR Pfam; PF00805; Pentapeptide; 2.  
KW Complete proteome.  
SQ SEQUENCE 584 AA; 65151 MW; 21BF5D5F0486CCC6 CRC64;

Query Match 56.3%; Score 40; DB 17; Length 584;  
Best Local Similarity 41.7%; Pred. No. 54;  
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Y 1 CFQWKRNMKVR 12  
| | | | | | | | | |  
b 445 CFTRKDMERK 456

RESULT 15  
J31090 PRELIMINARY; PRT; 273 AA.  
ID Q31090  
AC Q31090;

DT 01-JAN-1998 (TREMblrel. 05, Created)  
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)  
DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)  
DB Hypothetical 31.0 kDa protein.  
OS Rhizobium leguminosarum (biovar viciae).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Rhizobium.  
OX NCBI\_TaxID=387;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VF39;  
RA Ivashina T.V., Sadykov M.R., Kanapin A.A., Ksenzenko V.N.;  
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VF39;  
RX MEDLINE=99113394; PubMed=9914965;  
RA Sadykov M.R., Ivashina T.V., Kanapin A.A., Shlyapnikov M.G.,  
RA Ksenzenko V.N.;  
RT "Structural and functional organization of the exopolysaccharide  
biosynthesis genes in Rhizobium leguminosarum bv. viciae VF39.";  
RL Mol. Biol. (Mosk) 32:797-804(1998).  
DR EMBL; AF028810; AAB8891.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 273 AA; 30986 MW; F195D2FEF7AD44D3 CRC64;

Query Match 54.9%; Score 39; DB 2; Length 273;  
Best Local Similarity 60.0%; Pred. No. 37;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 QWKRNMKVR 12  
| | | | | | | | | |  
Db 245 RWLRNLRKLR 254

Search completed: February 21, 2003, 08:00:45  
Job time : 22.6744 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

DM protein - protein search, using sw model

Run on: February 21, 2003, 07:37:21 ; Search time 28.093 Seconds  
(without alignments)  
56.918 Million cell updates/sec

Title: US-09-743-107B-93

Perfect score: 69

Sequence: 1 CFQWKARKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 101002.\*  
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

result No.	Score	Query Match	Length	ID	Description
1	69	100.0	12	21	AAV78093 Human lactoferrin
2	65	94.2	12	21	AAV78074 Human lactoferrin
3	63	91.3	12	21	AAV78091 Human lactoferrin
4	63	91.3	12	21	AAV78092 Human lactoferrin
5	61	88.4	12	21	AAV78089 Human lactoferrin
6	61	88.4	12	21	AAV78090 Human lactoferrin
7	60	87.0	12	21	AAV78086 Human lactoferrin
8	59	85.5	12	21	AAV78038 Human lactoferrin
9	59	85.5	12	21	AAV78046 Human lactoferrin
10	59	85.5	12	21	AAV78047 Human lactoferrin

11	59	85.5	12	21	AAV78084 Human lactoferrin
12	59	85.5	12	21	AAV78095 Human lactoferrin
13	59	85.5	13	21	AAV78037 Human lactoferrin
14	59	85.5	13	21	AAV78048 Human lactoferrin
15	59	85.5	13	21	AAV78049 Human lactoferrin
16	59	85.5	14	21	AAV78036 Human lactoferrin
17	59	85.5	14	21	AAV78050 Human lactoferrin
18	59	85.5	14	21	AAV78051 Human lactoferrin
19	59	85.5	15	17	AAV78054 Peptide for anti-u
20	59	85.5	15	21	AAV78035 Human lactoferrin
21	59	85.5	15	21	AAV78062 Human lactoferrin
22	59	85.5	15	21	AAV78063 Human lactoferrin
23	59	85.5	16	21	AAV78031 Human lactoferrin
24	59	85.5	16	21	AAV78064 Human lactoferrin
25	59	85.5	16	21	AAV78065 Human lactoferrin
26	59	85.5	17	21	AAV78034 Human lactoferrin
27	59	85.5	17	21	AAV78066 Human lactoferrin
28	59	85.5	17	21	AAV78067 Human lactoferrin
29	59	85.5	18	15	AAV78032 Human lactoferrin
30	59	85.5	18	17	AAV78033 Human lactoferrin
31	59	85.5	18	21	AAV78033 Human lactoferrin
32	59	85.5	19	21	AAV78032 Human lactoferrin
33	59	85.5	19	21	AAV78032 Human lactoferrin
34	59	85.5	20	13	AAV78032 Human lactoferrin
35	59	85.5	20	14	AAV78032 Human lactoferrin
36	59	85.5	20	15	AAV78032 Human lactoferrin
37	59	85.5	20	15	AAV78032 Human lactoferrin
38	59	85.5	20	15	AAV78032 Human lactoferrin
39	59	85.5	20	15	AAV78032 Human lactoferrin
40	59	85.5	20	16	AAV78032 Human lactoferrin
41	59	85.5	20	16	AAV78032 Human lactoferrin
42	59	85.5	20	16	AAV78032 Human lactoferrin
43	59	85.5	20	16	AAV78032 Human lactoferrin
44	59	85.5	20	16	AAV78032 Human lactoferrin
45	59	85.5	20	17	AAV78032 Human lactoferrin

#### ALIGNMENTS

RESULT 1  
AAV78093  
ID AAV78093 standard; Peptide; 12 AA.  
AC AAV78093;  
XX 25-APR-2000 (first entry)  
DT Human lactoferrin derived peptide SEQ ID NO:93.  
DE Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW Bactericidal; preservative.  
OS Homo sapiens.  
OS Synthetic.  
XX WO200001730-A1.  
XX 13-JAN-2000.  
XX 06-JUL-1999; 99WO-SE01230.  
XX 06-JUL-1998; 98SE-0002441.  
XX 17-JUL-1998; 98SE-0002562.  
XX 29-DEC-1998; 98SE-0004614.  
XX (ASCI-) A+ SCI INVEST AB.  
XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX WPI; 2000-147388/13.



XX New peptides used for treatment and prevention of infections, -  
PT Inflammations and tumors and for use in infant formula food -  
XX  
XX Claim 22; Page 38; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.

XX  
XX SQ Sequence 12 AA;  
Query Match 100.0%; Score 69; DB 21; Length 12;  
Best Local Similarity 100.0%; Pred. No. 4.2e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CFQWKRAMRKVR 12  
| | | | | | | | | | | |  
Db 1 CFQWKRAMRKVR 12

RESULT 2  
AAY78074  
ID AAY78074 standard; Peptide; 12 AA.  
XX  
AC AAY78074;  
XX  
DT 25-APR-2000 (first entry)  
XX  
DE Human lactoferrin derived peptide SEQ ID NO:74.  
XX  
XX Human; lactoferrin; modification; infection; inflammation; tumour;  
XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
XX urinary tract infection; colitis; Candida infection; fungicidal;  
XX bactericidal; preservative.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
XX WO200001730-A1.  
XX  
XX 13-JAN-2000.  
XX  
XX 06-JUL-1999; 99WO-SB01230.  
XX  
XX 06-JUL-1998; 98SE-0002441.  
XX  
XX 17-JUL-1998; 98SE-0002562.  
XX  
XX 29-DEC-1998; 98SE-0004614.  
XX  
XX (ASCI-) A+ SCI INVEST AB.  
XX  
XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX  
XX WPI; 2000-147388/13.  
XX  
XX New peptides used for treatment and prevention of infections, -  
XX inflammations and tumors and for use in infant formula food -  
XX  
XX Claim 22; Page 35; 102pp; English.  
XX  
XX AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.

XX  
XX SQ Sequence 12 AA;  
Query Match 94.2%; Score 65; DB 21; Length 12;  
Best Local Similarity 91.7%; Pred. No. 0.00019;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CFQWKRAMRKVR 12  
| | | | | | | | | | | |  
Db 1 CFQWKRAMRKVR 12

RESULT 3  
AAY78091  
ID AAY78091 standard; Peptide; 12 AA.  
XX  
AC AAY78091;  
XX  
DT 25-APR-2000 (first entry)  
XX  
DE Human lactoferrin derived peptide SEQ ID NO:91.  
XX  
XX Human; lactoferrin; modification; infection; inflammation; tumour;  
XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
XX urinary tract infection; colitis; Candida infection; fungicidal;  
XX bactericidal; preservative.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
XX WO200001730-A1.  
XX  
XX 13-JAN-2000.  
XX  
XX 06-JUL-1999; 99WO-SB01230.  
XX  
XX 06-JUL-1998; 98SE-0002441.  
XX  
XX 17-JUL-1998; 98SE-0002562.  
XX  
XX 29-DEC-1998; 98SE-0004614.  
XX  
XX (ASCI-) A+ SCI INVEST AB.  
XX  
XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX  
XX WPI; 2000-147388/13.  
XX  
XX New peptides used for treatment and prevention of infections, -  
XX inflammations and tumors and for use in infant formula food -  
XX  
XX Claim 22; Page 38; 102pp; English.  
XX  
XX AAY78001 to AAY78100 represent peptides having sequences based on human  
XX lactoferrin. The peptides are taken up in the intestine through  
XX binding to specific lactoferrin receptors and are then transported  
XX through the circulation. A medicinal product of the peptide or fragment  
XX can be used for treating and/or prevention of infections (such as  
XX urinary tract infections, colitis, and Candida infection on a mucosal  
XX membrane), inflammations and/or tumours. The peptides can also be used  
XX in food stuffs such as infant formula food. The peptides are also

XX fungicidal and bactericidal and may also be used as preservatives.  
XX Even though native human lactoferrin have been shown to have desired  
XX anti-inflammatory anti-infectious and anti-tumoural properties they  
XX cannot be used clinically on a broad basis because of high production  
XX costs. Therefore, provision of peptides based on lactoferrin would  
XX enable them to be used for the same purposes as lactoferrin at lower  
XX cost.

QY Sequence 12 AA;

Query Match 91.3%; Score 63; DB 21; Length 12;

Best Local Similarity 91.7%; Pred. No. 0.00042; Mismatches 1; Conservative 0; Indels 0; Gaps 0;

QY 1 CFQWKRMKVR 12  
DB 1 CFQWKRMKVR 12

RESULT 4

AY78092  
D AAY78092 standard; Peptide; 12 AA.

XX AAY78092;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:92.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
XX urinary tract infection; colitis; Candida infection; fungicidal;  
XX bactericidal; preservative.

XX Homo sapiens.  
XX Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

XX 17-JUL-1998; 98SE-0002562.

XX 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI, 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
XX inflammations and tumors and for use in infant formula food -

XX Claim 22; Page 38; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
XX lactoferrin. The peptides are taken up in the intestine through  
XX binding to specific lactoferrin receptors and are then transported  
XX through the circulation. A medicinal product of the peptide or fragment  
XX can be used for treating and/or prevention of infections (such as  
XX urinary tract infections, colitis, and Candida infection on a mucosal  
XX membrane), inflammations and/or tumours. The peptides can also be used  
XX in food stuffs such as infant formula food. The peptides are also  
XX fungicidal and bactericidal and may also be used as preservatives.  
XX Even though native human lactoferrin have been shown to have desired  
XX anti-inflammatory anti-infectious and anti-tumoural properties they  
XX cannot be used clinically on a broad basis because of high production  
XX costs. Therefore, provision of peptides based on lactoferrin would  
XX enable them to be used for the same purposes as lactoferrin at lower  
XX cost.

XX SQ Sequence 12 AA;  
Query Match 91.3%; Score 63; DB 21; Length 12;  
Best Local Similarity 91.7%; Pred. No. 0.00042;  
Matches 1; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWKRMKVR 12  
DB 1 CFQWKRMKVR 12

RESULT 5

AY78089  
D AAY78089 standard; Peptide; 12 AA.

XX AAY78089;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:89.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
XX urinary tract infection; colitis; Candida infection; fungicidal;  
XX bactericidal; preservative.

XX Homo sapiens.  
XX Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

XX 17-JUL-1998; 98SE-0002562.

XX 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI, 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
XX inflammations and tumors and for use in infant formula food -

XX Claim 22; Page 37; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
XX lactoferrin. The peptides are taken up in the intestine through  
XX binding to specific lactoferrin receptors and are then transported  
XX through the circulation. A medicinal product of the peptide or fragment  
XX can be used for treating and/or prevention of infections (such as  
XX urinary tract infections, colitis, and Candida infection on a mucosal  
XX membrane), inflammations and/or tumours. The peptides can also be used  
XX in food stuffs such as infant formula food. The peptides are also  
XX fungicidal and bactericidal and may also be used as preservatives.  
XX Even though native human lactoferrin have been shown to have desired  
XX anti-inflammatory anti-infectious and anti-tumoural properties they  
XX cannot be used clinically on a broad basis because of high production  
XX costs. Therefore, provision of peptides based on lactoferrin would  
XX enable them to be used for the same purposes as lactoferrin at lower  
XX cost.

2y 1 CFQWXRMRKVR 12  
 ||||:|||||  
 3b 1 CFQWXRMRKVR 12

## RESULT 6

1AY78090  
 ID AAY78090 standard; Peptide; 12 AA.

XX AC AAY78090;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:90.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX OS Homo sapiens.  
 OS Synthetic.

XX PN WO200001730-A1.

XX FN 13-JAN-2000.

XX PD 06-JUL-1999; 99WO-S001230.

XX PF 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 PS Claim 22; Page 37; 102pp; English.

XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;

Query Match 88.4%; Score 61; DB 21; Length 12;  
 Best Local Similarity 83.3%; Pred. No. 0.0009;  
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRMRKVR 12

DB 1 CFQWXRMRKVR 12

## RESULT 7

AAY78086

ID AAY78086 standard; Peptide; 12 AA.

XX AC AAY78086;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:86.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX OS Homo sapiens.  
 OS Synthetic.

XX PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-S001230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 PS Claim 22; Page 36; 102pp; English.

XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;

Query Match 87.0%; Score 60; DB 21; Length 12;  
 Best Local Similarity 83.3%; Pred. No. 0.0013;  
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWXRMRKVR 12

DB 1 CFQWXRMRKVR 12

## RESULT 8

AAY78038

ID AAY78038 standard; Peptide; 12 AA.

XX AC AAY78038;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:38.

CX Human; lactoferrin; modification; infection; inflammation; tumour;  
 CW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 CW urinary tract infection; colitis; Candida infection; fungicidal;  
 CX bactericidal; preservative.  
 XS Homo sapiens.  
 XS Synthetic.  
 CX WO200001730-A1.  
 CX 13-JAN-2000.  
 CX 06-JUL-1999; 99WO-SE01230.  
 CX 06-JUL-1998; 98SE-0002441.  
 CX 17-JUL-1998; 98SE-0002562.  
 CX 29-DEC-1998; 98SE-0004614.  
 CX (ASCI-) A+ SCI INVEST AB.  
 CX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 CX WPI; 2000-147388/13.  
 CX New peptides used for treatment and prevention of infections,  
 CX inflammations and tumors and for use in infant formula food -  
 CX Claim 12; Page 70; 102pp; English.  
 CX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CX lactoferrin. The peptides are taken up in the intestine through  
 CX binding to specific lactoferrin receptors and are then transported  
 CX through the circulation. A medicinal product of the peptide or fragment  
 CX can be used for treating and/or prevention of infections (such as  
 CX urinary tract infections, colitis, and Candida infection on a mucosal  
 CX membrane), inflammations and/or tumours. The peptides can also be used  
 CX in food stuffs such as infant formula food. The peptides are also  
 CX fungicidal and bactericidal and may also be used as preservatives.  
 CX Even though native human lactoferrin have been shown to have desired  
 CX anti-inflammatory anti-infectious and anti-tumoural properties they  
 CX cannot be used clinically on a broad basis because of high production  
 CX costs. Therefore, provision of peptides based on lactoferrin would  
 CX enable them to be used for the same purposes as lactoferrin at lower  
 CX cost.  
 CX Sequence 12 AA;  
 CX Query Match 85.5%; Score 59; DB 21; Length 12;  
 CX Best Local Similarity 83.3%; Pred. No. 0.002; 1; Mismatches 0; Gaps 0;  
 CX Matches 10; Conservative 1; Indels 0; Gaps 0;  
 CX 1 CFQWKRAVRKVR 12  
 CX 1 CFQWQNRKVR 12  
 CX RESULT 9  
 CX AAY78046  
 CX ID AAY78046 standard; Peptide; 12 AA.  
 CX AC AAY78046;  
 CX DT 25-APR-2000 (first entry)  
 CX Human lactoferrin derived peptide SEQ ID NO:46.  
 CX Human; lactoferrin; modification; infection; inflammation; tumour;  
 CX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 CX urinary tract infection; colitis; Candida infection; fungicidal;  
 CX bactericidal; preservative.  
 CX Homo sapiens.

OS Synthetic.  
 XX WO200001730-A1.  
 XX 13-JAN-2000.  
 XX 06-JUL-1999; 99WO-SE01230.  
 XX 06-JUL-1998; 98SE-0002441.  
 XX 17-JUL-1998; 98SE-0002562.  
 XX 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 XX inflammations and tumors and for use in infant formula food -  
 XX Claim 15; Page 35; 102pp; English.  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 XX lactoferrin. The peptides are taken up in the intestine through  
 XX binding to specific lactoferrin receptors and are then transported  
 XX through the circulation. A medicinal product of the peptide or fragment  
 XX can be used for treating and/or prevention of infections (such as  
 XX urinary tract infections, colitis, and Candida infection on a mucosal  
 XX membrane), inflammations and/or tumours. The peptides can also be used  
 XX in food stuffs such as infant formula food. The peptides are also  
 XX fungicidal and bactericidal and may also be used as preservatives.  
 XX Even though native human lactoferrin have been shown to have desired  
 XX anti-inflammatory anti-infectious and anti-tumoural properties they  
 XX cannot be used clinically on a broad basis because of high production  
 XX costs. Therefore, provision of peptides based on lactoferrin would  
 XX enable them to be used for the same purposes as lactoferrin at lower  
 XX cost.  
 XX Sequence 12 AA;  
 XX Query Match 85.5%; Score 59; DB 21; Length 12;  
 XX Best Local Similarity 83.3%; Pred. No. 0.002; 1; Mismatches 0; Gaps 0;  
 XX Matches 10; Conservative 1; Indels 0; Gaps 0;  
 XX 1 CFQWKRAVRKVR 12  
 XX 1 CFQWQNRKVR 12  
 XX RESULT 10  
 XX AAY78047  
 XX ID AAY78047 standard; Peptide; 12 AA.  
 XX AC AAY78047;  
 XX DT 25-APR-2000 (first entry)  
 XX Human lactoferrin derived peptide SEQ ID NO:47.  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX urinary tract infection; colitis; Candida infection; fungicidal;  
 XX bactericidal; preservative.  
 XX Homo sapiens.  
 XX Synthetic.  
 XX WO200001730-A1.  
 XX 13-JAN-2000.  
 XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 18; Page 73; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 12 AA;

Query Match 85.5%; Score 59; DB 21; Length 12;  
 Best Local Similarity 83.3%; Pred. No. 0.002; 1; Indels 0; Gaps 0;  
 Matches 10; Conservative 1; Mismatches 1;

QY 1 CFQWKRAMRKVR 12  
 DB 1 CFQWQRNMRKVR 12

RESULT 11  
 AAY78084  
 ID AAY78084 standard; Peptide; 12 AA.

XX AAY78084;  
 AC 25-APR-2000 (first entry)  
 DT Human lactoferrin derived peptide SEQ ID NO:84.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.  
 OS Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 22; Page 36; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 12 AA;

Query Match 85.5%; Score 59; DB 21; Length 12;  
 Best Local Similarity 83.3%; Pred. No. 0.002; 1; Indels 0; Gaps 0;  
 Matches 10; Conservative 1; Mismatches 1;

QY 1 CFQWKRAMRKVR 12  
 DB 1 CFQWQRNMRKVR 12

RESULT 12  
 AAY78095  
 ID AAY78095 standard; Peptide; 12 AA.

XX AAY78095;

AC 25-APR-2000 (first entry)

DT Human lactoferrin derived peptide SEQ ID NO:95.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.

OS Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX

Claim 22; Page 38; 102pp; English.

AA78001 to AA78100 represent peptides having sequences based on human lactoferrin. The peptides are taken up in the intestine through binding to specific lactoferrin receptors and are then transported through the circulation. A medicinal product of the peptide or fragment can be used for treating and/or prevention of infections (such as urinary tract infections, colitis, and Candida infection) on a mucosal membrane), inflammations and/or tumours. The peptides can also be used in food stuffs such as infant formula food. The peptides are also fungicidal and bactericidal and may also be used as preservatives. Even though native human lactoferrin have been shown to have desired anti-inflammatory, anti-infectious and anti-tumoural properties they cannot be used clinically on a broad basis because of high production costs. Therefore, provision of peptides based on lactoferrin would enable them to be used for the same purposes as lactoferrin at lower cost.

Sequence 12 AA;

Query Match 95.5%; Score 59; DB 21; Length 12;  
Best Local Similarity 83.3%; Pred. No. 0.002;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Y 1 CFQWKRMRKVR 12  
| | | | | | | | | |  
b 1 CFQWKRMRKVR 12

RESULT 13

AA78037  
D AA78037 standard; Peptide, 13 AA.

C AA78037;

X 25-APR-2000 (first entry)

T Human lactoferrin derived peptide SEQ ID NO:37.

E Human; lactoferrin; modification; infection; inflammation; tumour;  
W food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
W urinary tract infection; colitis; Candida infection; fungicidal;  
W bactericidal; preservative.

X Homo sapiens.  
S Synthetic.

X WO200001730-A1.

N 13-JAN-2000.

D 06-JUL-1999; 99WO-SE01230.

F 06-JUL-1998; 98SE-0002441.

R 17-JUL-1998; 98SE-0002562.

R 29-DEC-1998; 98SE-0004614.

X (ASCI-) A+ SCI INVEST AB.

X Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

X WPI; 2000-147388/13.

X New peptides used for treatment and prevention of infections,  
T inflammations and tumors and for use in infant formula food -

S Claim 12; Page 70; 102pp; English.

AA78001 to AA78100 represent peptides having sequences based on human lactoferrin. The peptides are taken up in the intestine through binding to specific lactoferrin receptors and are then transported through the circulation. A medicinal product of the peptide or fragment can be used for treating and/or prevention of infections (such as

CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory, anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.

XX SQ Sequence 13 AA;

Query Match 85.5%; Score 59; DB 21; Length 13;  
Best Local Similarity 83.3%; Pred. No. 0.0021;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWKRMRKVR 12  
| | | | | | | | | |  
Db 2 CFQWKRMRKVR 13

RESULT 14

AA78048

ID AA78048 standard; Peptide, 13 AA.

XX AA78048;

XX 25-APR-2000 (first entry)

DT Human lactoferrin derived peptide SEQ ID NO:48.

DE Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.

XX Homo sapiens.  
OS Synthetic.

XX WO200001730-A1.

PN 13-JAN-2000.

PD 06-JUL-1999; 99WO-SE01230.

PF 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -

XX Claim 15; Page 74; 102pp; English.

AA78001 to AA78100 represent peptides having sequences based on human lactoferrin. The peptides are taken up in the intestine through binding to specific lactoferrin receptors and are then transported through the circulation. A medicinal product of the peptide or fragment can be used for treating and/or prevention of infections (such as urinary tract infections, colitis, and Candida infection) on a mucosal membrane), inflammations and/or tumours. The peptides can also be used in food stuffs such as infant formula food. The peptides are also fungicidal and bactericidal and may also be used as preservatives. Even though native human lactoferrin have been shown to have desired anti-inflammatory, anti-infectious and anti-tumoural properties they cannot be used clinically on a broad basis because of high production

CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.

XX Sequence 13 AA;

Query Match 85.5%; Score 59; DB 21; Length 13;  
Best Local Similarity 83.3%; Pred. No. 0.0021;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CFQWKRAMRKVR 12  
DB 2 CFQWQRNMRKVR 13

RESULT 15  
AA78049  
ID AA78049 standard; Peptide; 13 AA.

AC AA78049;  
DT 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:49.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.

XX Homo sapiens.  
XX Synthetic.

PN WO200001730-A1.

PD 13-JAN-2000.

PF 06-JUL-1999; 99WO-SE01230.

PR 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

PA (ASCI-) A+ SCI INVEST AB.

PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
XX inflammations and tumors and for use in infant formula food -  
XX Claim 18; Page 74; 102pp; English.

XX AA78001 to AA78100 represent peptides having sequences based on human  
XX lactoferrin. The peptides are taken up in the intestine through  
XX binding to specific lactoferrin receptors and are then transported  
XX through the circulation. A medicinal product of the peptide or fragment  
XX can be used for treating and/or prevention of infections (such as  
XX urinary tract infections, colitis, and Candida infection on a mucosal  
XX membrane), inflammations and/or tumours. The peptides can also be used  
XX in food stuffs such as infant formula food. The peptides are also  
XX fungicidal and bactericidal and may also be used as preservatives.  
XX Even though native human lactoferrin have been shown to have desired  
XX anti-inflammatory anti-infectious and anti-tumoural properties they  
XX cannot be used clinically on a broad basis because of high production  
XX costs. Therefore, provision of peptides based on lactoferrin would  
XX enable them to be used for the same purposes as lactoferrin at lower  
XX cost.

XX Sequence 13 AA;

Query Match 85.5%; Score 59; DB 21; Length 13;

Best Local Similarity 83.3%; Pred. No. 0.0021;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CFQWKRAMRKVR 12  
DB 2 CFQWQRNMRKVR 13

Search completed: February 21, 2003, 07:56:44  
Job time : 28.093 secs

GenCore version 5.1.3  
Copyright (C) 1993 - 2003 CompuGen Ltd.

M protein - protein search, using sw model

un on: February 21, 2003, 07:50:40 ; Search time 8.93023 Seconds  
(without alignments)  
39.537 Million cell updates/sec

title: US-09-743-107B-93  
effect score: 69  
sequence: 1 CFQWRAMRKVR 12

coring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

searched: 262574 seqs, 2942292 residues

total number of hits satisfying chosen parameters: 262574

inimum DB seq length: 0  
aximum DB seq length: 200000000

set-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

atadase : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A COMB.pap:\*  
2: /cgn2\_6/ptodata/1/1aa/5B COMB.pap:\*  
3: /cgn2\_6/ptodata/1/1aa/6A COMB.pap:\*  
4: /cgn2\_6/ptodata/1/1aa/6B COMB.pap:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS COMB.pap:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pap:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

result	No.	Score	Query Match	Length	DB ID	Description
1	59	85.5	18	1	US-08-204-487-3	Sequence 3, Appli
2	59	85.5	18	2	US-08-485-948-8	Sequence 8, Appli
3	59	85.5	18	2	US-08-628-380-8	Sequence 8, Appli
4	59	85.5	18	2	US-08-475-055-8	Sequence 8, Appli
5	59	85.5	20	1	US-07-753-161A-3	Sequence 3, Appli
6	59	85.5	20	1	US-07-891-174-3	Sequence 3, Appli
7	59	85.5	20	1	US-08-204-487-1	Sequence 1, Appli
8	59	85.5	20	1	US-08-256-771-24	Sequence 24, Appli
9	59	85.5	20	1	US-08-256-771-25	Sequence 24, Appli
10	59	85.5	20	1	US-08-381-984-24	Sequence 24, Appli
11	59	85.5	20	1	US-08-381-984-25	Sequence 25, Appli
12	59	85.5	22	4	US-09-508-734-4	Sequence 4, Appli
13	59	85.5	24	4	US-09-508-734-6	Sequence 4, Appli
14	59	85.5	25	1	US-07-753-161A-10	Sequence 10, Appl
15	59	85.5	25	1	US-07-891-174-10	Sequence 10, Appl
16	59	85.5	25	1	US-08-204-487-7	Sequence 7, Appli
17	59	85.5	29	4	US-09-508-734-8	Sequence 8, Appli
18	59	85.5	36	1	US-07-753-161A-8	Sequence 8, Appli
19	59	85.5	36	1	US-07-891-174-8	Sequence 8, Appli
20	59	85.5	36	1	US-08-256-771-30	Sequence 30, Appl
21	59	85.5	36	1	US-08-381-984-29	Sequence 29, Appl
22	59	85.5	47	2	US-08-464-182A-6	Sequence 6, Appli
23	59	85.5	47	2	US-08-406-271-6	Sequence 6, Appli
24	59	85.5	50	2	US-08-693-274A-7	Sequence 7, Appli
25	59	85.5	52	4	US-09-017-043A-3	Sequence 3, Appli
26	59	85.5	53	2	US-08-464-182A-5	Sequence 5, Appli
27	59	85.5	53	2	US-08-406-271-5	Sequence 5, Appli

28	59	85.5	54	2	US-08-464-182A-2	Sequence 2, Appli
29	59	85.5	54	2	US-08-406-271-2	Sequence 2, Appli
30	59	85.5	694	3	US-08-724-586-2	Sequence 2, Appli
31	59	85.5	694	4	US-09-431-632-2	Sequence 2, Appli
32	59	85.5	694	4	US-09-932-190-2	Sequence 2, Appli
33	59	85.5	705	2	US-08-685-640-2	Sequence 2, Appli
34	59	85.5	708	2	US-08-685-640-4	Sequence 4, Appli
35	59	85.5	711	1	US-08-154-019-4	Sequence 4, Appli
36	59	85.5	711	1	US-08-461-333-4	Sequence 4, Appli
37	59	85.5	711	3	US-08-464-167-4	Sequence 4, Appli
38	59	85.5	711	3	US-09-158-313-4	Sequence 4, Appli
39	59	85.5	711	4	US-08-476-798-4	Sequence 4, Appli
40	56	81.2	711	1	US-08-145-681-2	Sequence 2, Appli
41	56	81.2	711	1	US-08-250-308-2	Sequence 2, Appli
42	56	81.2	711	1	US-08-453-703-2	Sequence 2, Appli
43	56	81.2	711	2	US-08-456-106-2	Sequence 2, Appli
44	56	81.2	711	3	US-08-456-108-2	Sequence 2, Appli
45	56	81.2	711	4	US-09-265-577-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-08-204-487-3  
; Sequence 3, Application US/08204487  
; Patent No. 5565425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAOKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGERAKI  
; APPLICANT: DOSAKO, SHUN'ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; TITLE OF INVENTION: INHIBITORS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: THIÉBAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FUN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..18  
; OTHER INFORMATION: /note= "HUMAN LACTOFERRIN PEPTIDE"



OTHER INFORMATION: (20-37)"

Query Match 85.5%; Score 59; DB 1; Length 18;  
Best Local Similarity 83.3%; Pred. No. 0.0015;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Dy 1 CFQWKRAMRKVR 12  
Db 1 CFQWQNMNRKVR 12

RESULT 2

US-08-485-948-8  
; Sequence 8, Application US/08485948  
; Patent No. 5855882  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VIASSARA  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,948  
; FILING DATE:  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/488,217  
; FILING DATE: JUNE 7, 1995  
; APPLICATION NUMBER: 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LP-Cl, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
US-08-485-948-8

Query Match 85.5%; Score 59; DB 2; Length 18;  
Best Local Similarity 83.3%; Pred. No. 0.0015;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Dy 1 CFQWKRAMRKVR 12  
Db 1 CFQWQNMNRKVR 12

RESULT 3

US-08-628-380-8  
; Sequence 8, Application US/08628380  
; Patent No. 5891341  
; GENERAL INFORMATION:  
; APPLICANT: LI, YONG MING  
; APPLICANT: VIASSARA, HELEN  
; APPLICANT: CERAMI, ANTHONY  
; TITLE OF INVENTION: AGE-MEDIATED INHIBITION OF ANTIBACTERIAL PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/628,380  
; FILING DATE: April 4, 1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LP-Cl, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
US-08-628-380-8

Query Match 85.5%; Score 59; DB 2; Length 18;  
Best Local Similarity 83.3%; Pred. No. 0.0015;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Dy 1 CFQWKRAMRKVR 12  
Db 1 CFQWQNMNRKVR 12

RESULT 4

US-08-475-055-8  
; Sequence 8, Application US/08475055  
; Patent No. 5962245  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VIASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/475,055  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/485,948  
FILING DATE:  
APPLICATION NUMBER: 08/488,217  
FILING DATE: JUNE 7, 1995  
APPLICATION NUMBER: 08/418,642  
FILING DATE: APRIL 7, 1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 947-1-008A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
DESCRIPTION: LP-CL, 8-25  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
S-08-475-055-8

Query Match 85.5%; Score 59; DB 2; Length 18;  
Best Local Similarity 83.3%; Pred. No. 0.0015;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

y 1 CFQWXRMRKVR 12  
|||:|||||  
b 1 CFQWQRMRKVR 12

RESULT 5  
S-07-755-161A-3  
Sequence 3, Application US/07755161A  
Patent No: 5304633  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: Antimicrobial Peptides and an  
TITLE OF INVENTION: Antimicrobial Agent  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/755.161A  
FILING DATE: 19910905  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX: 202-371-8856  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 2  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 19"  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 19  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 2"  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLES:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-755-161A-3

Query Match 85.5%; Score 59; DB 1; Length 20;

Best Local Similarity 83.3%; Pred. No. 0.0016; Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 CFOWKRAMRKVR 12  
2 CFOWQRMNRKVR 13

RESULT 6  
JS-07-891-174-3  
Sequence 3, Application US/07891174  
Patent No. 5317084  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: Antimicrobial Peptides and an  
TITLE OF INVENTION: Antimicrobial Agent  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Displaywrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/891,174  
FILING DATE: 29-MAY-1992  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/755,161  
FILING DATE: 05-SEP-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX: 202-371-8856  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY: modified site

LOCATION: 2  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 19"  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 19  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 2"  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-891-174-3

Query Match 85.5%; Score 59; DB 1; Length 20;  
Best Local Similarity 83.3%; Pred. No. 0.0016; Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 CFOWKRAMRKVR 12  
2 CFOWQRMNRKVR 13

RESULT 7  
US-08-204-487-1  
Sequence 1, Application US/08204487  
Patent No. 5565425  
GENERAL INFORMATION:  
APPLICANT: YAMAMOTO, NAOKI  
APPLICANT: NAKASHIMA, HIDEKI  
APPLICANT: MOSUCHI, WATARU  
APPLICANT: TANAKA, SHIGEAKI  
APPLICANT: DOSAKO, SHUN-ICHI  
APPLICANT: KAWASAKI, YOSHIHIRO  
APPLICANT: UCHIDA, TOSHIAKI  
TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
TITLE OF INVENTION: INHIBITORS  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
ADDRESS: THIBEAULT  
STREET: 53 STATE STREET  
CITY: BOSTON  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/204,487  
FILING DATE: 02-MAR-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: CAMPBELL, PAULA A.  
REGISTRATION NUMBER: 32,503  
REFERENCE/DOCKET NUMBER: FJN-019  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..20  
OTHER INFORMATION: /note= "ANTIBACTERIAL PEPTIDE  
DERIVED FROM HUMAN LACTOFERRIN"  
S-08-204-487-1

Query Match 85.5%; Score 59; DB 1; Length 20;  
Best Local Similarity 83.3%; Pred. No. 0.0016; Indels 0; Gaps 0;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Y 1 CFQWKRAVRKVR 12  
| | | | |  
b 2 CFQWQRNVRKVR 13  
| | | | |

RESULT 8  
S-08-256-771-24  
Sequence 24, Application US/08256771  
Patent No. 5656591  
GENERAL INFORMATION:  
APPLICANT: Manoru TOMITA et al.  
TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
TITLE OF INVENTION: PRODUCTS THEREWITH  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,771  
FILING DATE: July 22, 1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "Cys residues are linked by  
disulfide bond"

US-08-256-771-24

Query Match 85.5%; Score 59; DB 1; Length 20;  
Best Local Similarity 83.3%; Pred. No. 0.0016;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CFQWKRAVRKVR 12  
| | | | |  
DB 2 CFQWQRNVRKVR 13  
| | | | |

RESULT 9  
US-08-256-771-25  
Sequence 25, Application US/08256771  
Patent No. 5656591  
GENERAL INFORMATION:  
APPLICANT: Manoru TOMITA et al.  
TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
TITLE OF INVENTION: PRODUCTS THEREWITH  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,771  
FILING DATE: July 22, 1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "Cys residues are protected to  
prevent disulfide bond"  
US-08-256-771-25

Query Match 85.5%; Score 59; DB 1; Length 20;  
Best Local Similarity 83.3%; Pred. No. 0.0016;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CFQWKRAVRKVR 12  
| | | | |  
DB 2 CFQWQRNVRKVR 13  
| | | | |

RESULT 10  
US-08-381-984-24

Sequence 24, Application US/08381984  
Patent No 5804555  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIOXIDANT  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/381,984  
FILING DATE: April 11, 1995  
CLASSIFICATION: 252  
PRIOR APPLICATION NUMBER:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:

## INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "cysteine residues at positions 2  
OTHER INFORMATION: and 19 are bonded by disulfide linkage"

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "the specified peptide as well as  
OTHER INFORMATION: peptides including the specified peptide as a fragment thereof"

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "the specified peptide as well as  
OTHER INFORMATION: peptides including the specified peptide as a fragment thereof"

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "the specified peptide as well as  
OTHER INFORMATION: peptides including the specified peptide as a fragment thereof"

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "the specified peptide as well as  
OTHER INFORMATION: peptides including the specified peptide as a fragment thereof"

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "the specified peptide as well as  
OTHER INFORMATION: peptides including the specified peptide as a fragment thereof"

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "the specified peptide as well as  
OTHER INFORMATION: peptides including the specified peptide as a fragment thereof"

FEATURE:

STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/381,984  
FILING DATE: April 11, 1995  
CLASSIFICATION: 252  
PRIOR APPLICATION NUMBER:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:

## INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "the specified peptide as well as  
OTHER INFORMATION: peptides including the specified peptide as a fragment thereof"

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "cysteine residues at positions 2  
OTHER INFORMATION: and 19 are chemically modified to prevent disulfide linkage"

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "cysteine residues at positions 2  
OTHER INFORMATION: and 19 are chemically modified to prevent disulfide linkage"

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "cysteine residues at positions 2  
OTHER INFORMATION: and 19 are chemically modified to prevent disulfide linkage"

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "cysteine residues at positions 2  
OTHER INFORMATION: and 19 are chemically modified to prevent disulfide linkage"

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "cysteine residues at positions 2  
OTHER INFORMATION: and 19 are chemically modified to prevent disulfide linkage"

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "cysteine residues at positions 2  
OTHER INFORMATION: and 19 are chemically modified to prevent disulfide linkage"

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "cysteine residues at positions 2  
OTHER INFORMATION: and 19 are chemically modified to prevent disulfide linkage"

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "cysteine residues at positions 2  
OTHER INFORMATION: and 19 are chemically modified to prevent disulfide linkage"

```
LENGTH: 22
TYPE: PRT
ORGANISM: Homo sapiens
3-09-508-734-4

Query Match      85.5%; Score 59; DB 4; Length 22;
Best Local Similarity 83.3%; Pred. No. 0.0018;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Y 1 CFQWKREMRKVR 12
   |||||
2 2 CFQWQRNMRKVR 13

RESULT 13
3-09-508-734-6
Sequence 6, Application US/09508734
Patent No. 6423509
GENERAL INFORMATION:
APPLICANT: Samyang Genex Corporation
TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and
FILE REFERENCE: PA/SYG/00139
CURRENT APPLICATION NUMBER: US/09/508,734
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/KR99/00373
PRIOR FILING DATE: 1998-07-14
PRIOR APPLICATION NUMBER: KR1998-29351
PRIOR FILING DATE: 1998-07-13
NUMBER OF SEQ ID NOS: 12
SOFTWARE: KopatentIn 1.71
SEQ ID NO 6
LENGTH: 24
TYPE: PRT
ORGANISM: Homo sapiens
3-09-508-734-6

Query Match      85.5%; Score 59; DB 4; Length 24;
Best Local Similarity 83.3%; Pred. No. 0.0019;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Y 1 CFQWKREMRKVR 12
   |||||
2 3 CFQWQRNMRKVR 14

RESULT 14
3-07-755-161A-10
Sequence 10, Application US/07755161A
Patent No. 5304633
GENERAL INFORMATION:
APPLICANT: Mamoru TOMITA et al.
TITLE OF INVENTION: Antimicrobial Peptides and an
OPERATING SYSTEM: MS-DOS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSER: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/755,161A
FILING DATE: 19910905
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
```

```
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX: 202-371-8856
TELEX:
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY: modified site
LOCATION: 4
IDENTIFICATION METHOD:
OTHER INFORMATION: /note= "thiol group of
OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with
OTHER INFORMATION: thiol group of Cys residue at location 21"
FEATURE:
NAME/KEY: modified site
LOCATION: 21
IDENTIFICATION METHOD:
OTHER INFORMATION: /note= "thiol group of
OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with
OTHER INFORMATION: thiol group of Cys residue at location 4"
PUBLICATION INFORMATION:
AUTHORS:
TITLES:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-07-755-161A-10

Query Match      85.5%; Score 59; DB 1; Length 25;
Best Local Similarity 83.3%; Pred. No. 0.002;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWKREMRKVR 12
   |||||
DB 4 CFQWQRNMRKVR 15
```

## RESULT 15

US-07-891-174-10  
Sequence 10, Application US/07891174  
Patent No. 5317084  
GENERAL INFORMATION:  
APPLICANT: MAMOTU TOMITA et al.  
TITLE OF INVENTION: Antimicrobial peptides and an  
TITLE OF INVENTION: Antimicrobial Agent  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/891,174  
FILING DATE: 29-MAY-1992  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/755,161  
FILING DATE: 05-SEP-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX: 202-371-8856  
TELEX:  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 4  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 21"  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 21

IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 4"  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-891-174-10  
Query Match 85.5%; Score 59; DB 1; Length 25;  
Best Local Similarity 83.3%; Pred. No. 0.002;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CFQWGRAMRKVR 12  
DB 4 CFQWGRAMRKVR 15  
Search completed: February 21, 2003, 08:04:27  
Job time : 9.93023 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

M protein - protein search, using sw model

un on: February 21, 2003, 08:08:15 ; Search time 6.88372 Seconds  
(Without alignments)  
54.162 Million cell updates/sec

title: US-09-743-107b-93

effect score: 69

sequence: 1 CFQWKRMRKVR 12

coring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

searched: 156504 seqs, 31069816 residues

total number of hits satisfying chosen parameters: 156504

inimum DB seq length: 0

aximum DB seq length: 200000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

atabase :

Published Applications AA.\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PTCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US05\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/PTCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	59	85.5	15	9	US-09-798-869-2
2	59	85.5	25	9	US-09-798-869-20
3	59	85.5	694	9	US-10-023-096-2
4	51	73.9	15	9	US-09-798-869-6
5	49	71.0	15	9	US-09-798-869-3
6	49	71.0	25	9	US-09-798-869-23
7	41	59.4	15	9	US-09-798-869-7
8	40	58.0	15	9	US-09-798-869-4
9	40	58.0	25	9	US-09-798-869-22
10	38	55.1	21	10	US-09-864-761-47985
11	37	53.6	15	9	US-09-798-869-8
12	37	53.6	15	9	US-09-798-869-29
13	37	53.6	15	9	US-09-798-869-30
14	37	53.6	489	9	US-09-888-320-2
15	36	52.2	86	9	US-09-738-626-5715
16	36	52.2	318	12	US-10-042-417-42
17	35	50.7	301	9	US-10-080-960-11
18	35	50.7	302	10	US-09-948-078-2
19	35	50.7	333	9	US-10-001-054-42

20	35	50.7	333	9	US-10-028-072-132	Sequence 132, App
21	35	50.7	333	9	US-10-121-049-132	Sequence 132, App
22	35	50.7	333	9	US-10-123-904-132	Sequence 132, App
23	35	50.7	333	9	US-10-140-470-132	Sequence 132, App
24	35	50.7	333	9	US-10-175-746-132	Sequence 132, App
25	35	50.7	333	9	US-10-176-918-132	Sequence 132, App
26	35	50.7	333	9	US-10-176-921-132	Sequence 132, App
27	35	50.7	333	9	US-10-137-865-132	Sequence 132, App
28	35	50.7	333	9	US-10-140-474-132	Sequence 132, App
29	35	50.7	351	10	US-09-853-625B-16	Sequence 16, Appl
30	34	49.3	138	9	US-10-013-379-14	Sequence 14, Appl
31	34	49.3	760	10	US-09-866-582-42	Sequence 42, Appl
32	33	47.8	95	10	US-09-764-864-1031	Sequence 1031, Ap
33	33	47.8	162	9	US-09-738-626-6604	Sequence 6604, Ap
34	33	47.8	238	9	US-09-986-480-182	Sequence 182, App
35	33	47.8	268	10	US-09-764-864-1172	Sequence 1172, Ap
36	33	47.8	411	10	US-09-764-864-899	Sequence 899, App
37	33	47.8	411	10	US-09-764-864-1346	Sequence 1346, Ap
38	33	47.8	414	9	US-09-978-295A-466	Sequence 466, App
39	33	47.8	414	9	US-09-764-857-12	Sequence 12, Appl
40	33	47.8	414	9	US-09-978-697-466	Sequence 466, App
41	33	47.8	414	9	US-09-978-192A-466	Sequence 466, App
42	33	47.8	414	9	US-09-998-832A-466	Sequence 466, App
43	33	47.8	414	9	US-09-978-189-466	Sequence 466, App
44	33	47.8	414	9	US-10-028-072-418	Sequence 418, App
45	33	47.8	414	9	US-10-121-049-418	Sequence 418, App

#### ALIGNMENTS

##### RESULT 1

US-09-798-869-2  
; Sequence 2, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798, 869  
; PRIOR FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-09-798-869-2

Query Match 85.5%; Score 59; DB 9; Length 15;  
Best Local Similarity 83.3%; Pred No. 0.0047;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQWKRMRKVR 12  
| | | | | | | | | |  
Db 3 CFQWKRMRKVR 14

##### RESULT 2

US-09-798-869-20  
; Sequence 20, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ RNSSON



APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 20  
LENGTH: 25  
TYPE: PRT  
ORGANISM: Homo Sapiens  
US-09-798-869-20

Query Match 85.5%; Score 59; DB 9; Length 25;  
Best Local Similarity 83.3%; Pred. No. 0.00078;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWKRAMRKVR 12  
|||:|||||  
DB 3 CFQWQNNRKVR 14

RESULT 3  
US-10-023-096-2  
Sequence 2, Application US/10023096  
Patent No. US20020160941A1  
GENERAL INFORMATION:  
APPLICANT: Krusel, Marian L.  
APPLICANT: Kurecki, Tomasz  
APPLICANT: Gollnick, Paul D.  
APPLICANT: Doyle, Darrell J.  
TITLE OF INVENTION: Cloning, Expression, and Uses of Human  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jacobson, Price, Holman & Stern  
STREET: 400 Seventh St. N.W.  
CITY: Washington D.C.  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/023,096  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,586  
FILING DATE: 30-SEPT-1996  
APPLICATION NUMBER: US 08/238,445  
FILING DATE: 05-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Player, William E.  
REGISTRATION NUMBER: 31,409  
REFERENCE/DOCKET NUMBER: 10505/P58185C  
TELEPHONE: (202) 638-6666  
TELEFAX: (202) 393-5350  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 694 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-10-023-096-2

Query Match 85.5%; Score 59; DB 9; Length 694;  
Best Local Similarity 83.3%; Pred. No. 0.019;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWKRAMRKVR 12  
|||:|||||  
DB 22 CFQWQNNRKVR 33

RESULT 4  
US-09-798-869-6  
Sequence 6, Application US/09798869  
Publication No. US20030022821A1  
GENERAL INFORMATION:  
APPLICANT: JOHN SIGURD SVENDSEN  
APPLICANT: (VSTEIN REKDAL  
APPLICANT: BALDUR SVEINBJ (RNSSON  
APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of Homo sapiens  
OTHER INFORMATION: sequence)  
US-09-798-869-6

Query Match 73.9%; Score 51; DB 9; Length 15;  
Best Local Similarity 75.0%; Pred. No. 0.01;  
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWKRAMRKVR 12  
|||:|||||  
DB 3 CFQWQNNRKVR 14

RESULT 5  
US-09-798-869-3  
Sequence 3, Application US/09798869  
Publication No. US20030022821A1  
GENERAL INFORMATION:  
APPLICANT: JOHN SIGURD SVENDSEN  
APPLICANT: (VSTEIN REKDAL  
APPLICANT: BALDUR SVEINBJ (RNSSON  
APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 15  
TYPE: PRT  
ORGANISM: CAPRINE  
US-09-798-869-3

Query Match 71.0%; Score 49; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.022;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 CFQWKRAMRKV 11  
|:|:|:|:|:  
3 CYQWQRMREKL 13

## RESULT 6

3-09-798-869-23  
Sequence 23, Application US/09798869  
Publication No. US20030022821A1  
GENERAL INFORMATION:  
APPLICANT: JOHN SIGURD SVENDSEN  
APPLICANT: (YSTEIN REKDAL  
APPLICANT: BALDUR SVEINBJ (RNSSON  
APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 23  
LENGTH: 25  
TYPE: PRT  
ORGANISM: CAPRINE  
3-09-798-869-23

Query Match 71.0%; Score 49; DB 9; Length 25;  
Best Local Similarity 63.6%; Pred. No. 0.037; Indels 1; Gaps 0;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 CFQWKRAMRKV 11  
|:|:|:|:|:  
3 CYQWQRMREKL 13

## RESULT 7

3-09-798-869-7  
Sequence 7, Application US/09798869  
Publication No. US20030022821A1  
GENERAL INFORMATION:  
APPLICANT: JOHN SIGURD SVENDSEN  
APPLICANT: (YSTEIN REKDAL  
APPLICANT: BALDUR SVEINBJ (RNSSON  
APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 7  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Artificial Sequence

OTHER INFORMATION: synthetic peptide (modified form of homo sapiens  
OTHER INFORMATION: sequence)

## 3-09-798-869-7

Query Match 59.4%; Score 41; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.49; Indels 2; Gaps 0;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWKRAMRKV 11  
|:|:|:|:|:  
DB 3 CYQWQRMREKL 13

## RESULT 8

US-09-798-869-4  
Sequence 4, Application US/09798869  
Publication No. US20030022821A1  
GENERAL INFORMATION:  
APPLICANT: JOHN SIGURD SVENDSEN  
APPLICANT: (YSTEIN REKDAL  
APPLICANT: BALDUR SVEINBJ (RNSSON  
APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 15  
TYPE: PRT  
ORGANISM: MURINE  
US-09-798-869-4

Query Match 58.0%; Score 40; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.71; Indels 3; Gaps 0;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWKRAMRKV 11  
|:|:|:|:|:  
DB 3 CLRWQNMREKV 13

## RESULT 9

US-09-798-869-22  
Sequence 22, Application US/09798869  
Publication No. US20030022821A1  
GENERAL INFORMATION:  
APPLICANT: JOHN SIGURD SVENDSEN  
APPLICANT: (YSTEIN REKDAL  
APPLICANT: BALDUR SVEINBJ (RNSSON  
APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 22  
LENGTH: 25  
TYPE: PRT  
ORGANISM: MURINE  
US-09-798-869-22

Query Match 58.0%; Score 40; DB 9; Length 25;  
Best Local Similarity 54.5%; Pred. No. 1.2; Indels 3; Gaps 0;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWKRAMRKV 11  
|:|:|:|:|:  
DB 3 CLRWQNMREKV 13

RESULT 10  
US-09-864-761-47985  
; Sequence 47985, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aconica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; PRIOR FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 47985  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL096701.14  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6  
; OTHER INFORMATION: EST HUMAN HIT: AW294800.1, EVALUATE 1.00e-06  
US-09-864-761-47985

Query Match 55.1%; Score 38; DB 10; Length 21;  
Best Local Similarity 83.3%; Pred. No. 2.1;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFQWKR 6  
| | | | |  
Db 16 CFQWRR 21

RESULT 11  
US-09-798-869-8  
; Sequence 8, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; PRIOR FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of caprine  
; OTHER INFORMATION: sequence)  
US-09-798-869-8

Query Match 53.6%; Score 37; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 2.3;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWKRAMRKV 11  
| | | | |  
Db 3 CLRQWEMRKV 13

RESULT 12  
US-09-798-869-29  
; Sequence 29, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; PRIOR FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 29  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: BOVINE  
US-09-798-869-29

Query Match 53.6%; Score 37; DB 9; Length 15;  
Best Local Similarity 45.5%; Pred. No. 2.3;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWKRAMRKV 11  
| | | | |  
Db 3 CFRQWEMKKL 13

RESULT 13  
US-09-798-869-30

Sequence 30, Application US/09798869  
Publication No. US2003002282A1  
GENERAL INFORMATION:

APPLICANT: JOHN SIGURD SVENDSEN  
APPLICANT: (YSTEIN REKDAL  
APPLICANT: BALDUR SVEINBJ (RNSSON  
APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 30  
LENGTH: 15

TYPE: PRT  
ORGANISM: BOVINE  
US-09-798-869-30

Query Match 53.6%; Score 37; DB 9; Length 15;  
Best Local Similarity 45.5%; Pred. No. 2.3;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

1 CFQWKRMKV 11  
|:|:|:|:|:  
3 CFQWQRMKL 13

RESULT 14  
US-09-888-320-2  
Sequence 2, Application US/09888320  
Publication No. US20030013090A1  
GENERAL INFORMATION:

APPLICANT: Barry III, Clifton E.  
APPLICANT: DeBarber, Andrea E.  
APPLICANT: Mdululi, Khisimuzi  
APPLICANT: Bekker, Linda-Gail  
APPLICANT: The Government of the United States of America  
APPLICANT: as represented by The Secretary of the  
APPLICANT: Department of Health and Human Services  
TITLE OF INVENTION: Methods of Diagnosing Multidrug Resistant Tuberculosis  
FILE REFERENCE: 015280-413100J3  
CURRENT APPLICATION NUMBER: US/09/888,320  
CURRENT FILING DATE: 2001-06-22  
PRIOR APPLICATION NUMBER: US 60/214,187  
PRIOR FILING DATE: 2000-06-26  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 489

TYPE: PRT  
ORGANISM: Mycobacterium tuberculosis  
OTHER INFORMATION: wild-type Etaa monooxygenase (RV3854C, Echa)  
US-09-888-320-2

Query Match 53.6%; Score 37; DB 9; Length 489;  
Best Local Similarity 54.5%; Pred. No. 66;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

1 CFQWKRMKV 11  
|:|:|:|:|:  
253 CQKPRRMRKM 263

RESULT 15  
US-09-738-626-5715  
Sequence 5715, Application US/09738626  
Publication No. US20020197605A1  
GENERAL INFORMATION:

APPLICANT: NAKAGAWA, SATOSHI  
APPLICANT: MIZOGUCHI, HIROSHI  
APPLICANT: ANDO, SEIKO  
APPLICANT: HAYASHI, MIKIRO  
APPLICANT: OCHIALI, KEIKO  
APPLICANT: YOKOI, HARUHIKO  
APPLICANT: TATEISHI, NAKO  
APPLICANT: SENOH, AKIHIRO  
APPLICANT: IKEDA, MASATO  
APPLICANT: OZAKI, AKIO  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-125  
CURRENT APPLICATION NUMBER: US/09/738,626  
CURRENT FILING DATE: 2000-12-18  
PRIOR APPLICATION NUMBER: JP 99/377484  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: JP 00/159162  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: JP 00/280988  
PRIOR FILING DATE: 2000-08-03  
NUMBER OF SEQ ID NOS: 7059  
SOFTWARE: PatentIn ver. 3.0  
SEQ ID NO 5715  
LENGTH: 86  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-738-626-5715

Query Match 52.2%; Score 36; DB 9; Length 86;  
Best Local Similarity 45.5%; Pred. No. 18;  
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 2 FQWKRMKV 12  
|:|:|:|:|:  
Db 73 FEYRQLRKIR 83

Search completed: February 21, 2003, 08:11:57  
Job time : 7.88372 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

1 protein - protein search, using sw model

in on: February 21, 2003, 07:48:01 ; Search time 10.6047 Seconds  
(without alignments)  
108.784 Million cell updates/sec

tle: US-09-743-107b-93  
rfect score: 69  
quence: 1 CFQWKRAVRKVR 12

oring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

arched: 283224 seqs, 96134422 residues  
al number of hits satisfying chosen parameters: 283224

nimum DB seq length: 0  
ximum DB seq length: 2000000000  
st-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

atabase : PIR 73:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	85.5	711	1	TFHUL
2	49	71.0	708	2	lactoferrin - goat
3	46	66.7	33	2	lactoferrin - shee
4	45	65.2	275	2	hypothetical prote
5	43	62.3	206	2	pyridoxamine 5'-ph
6	43	62.3	206	2	pyridoxamine 5'-ph
7	43	62.3	208	2	probable pyridoxam
8	42	60.9	1135	2	phytochrome C - so
9	41	59.4	1391	2	hypothetical prote
10	41	59.4	1397	2	protein F09C3.1 [i
11	40	58.0	707	1	lactoferrin precu
12	39	56.5	453	2	hypothetical prote
13	39	56.5	458	2	dyein beta heavy
14	38	55.1	275	2	32k protein - vacc
15	38	55.1	377	2	hypothetical prote
16	38	55.1	377	2	35k myristylprotei
17	38	55.1	377	2	A17L protein - var
18	38	55.1	378	2	A16L protein - var
19	38	55.1	500	2	ADP ATP carrier pr
20	38	55.1	2594	2	kinase-related pro
21	37	53.6	202	2	hypothetical prote
22	37	53.6	380	2	hypothetical prote
23	37	53.6	433	2	hypothetical prote
24	37	53.6	489	2	probable monooxyge
25	37	53.6	489	2	calnexin-like prot
26	37	53.6	530	2	calnexin homolog -
27	37	53.6	584	2	hypothetical prote
28	37	53.6	584	2	probable vacuolar
29	37	53.6	932	2	hypothetical prote

30 37 53.6 1131 2 T09701  
31 36 52.2 99 2 AC1714  
32 36 52.2 99 2 AG1343  
33 36 52.2 124 2 C96582  
34 36 52.2 158 2 D82782  
35 36 52.2 193 2 S01975  
36 36 52.2 289 2 G86403  
37 36 52.2 299 2 G95232  
38 36 52.2 303 1 WZBEM6  
39 36 52.2 303 2 C33374  
40 36 52.2 401 2 T41387  
41 36 52.2 428 2 T03251  
42 36 52.2 501 2 T04735  
43 36 52.2 502 2 T07141  
44 36 52.2 509 1 S62899  
45 36 52.2 510 2 T07119

ALIGNMENTS

RESULT 1

TFHUL  
lactotransferrin precursor [validated] - human  
N:Alternate names: lactoferrin  
C:Species: Homo sapiens (man)  
C:Date: 31-Mar-1992 #sequence, revision 21-Nov-1997 #text change 08-Dec-2000  
C:Accession: G01394; S11228; A45401; S10324; S15853; S20841; S07160; A61169; A31000; S743  
R:Cho, Y.  
submitted to the EMBL Data Library, March 1994  
A:Reference number: G06820  
A:Accession: G01394  
A:Status: preliminary; translated from GB/EMBL/DDSV  
A:Molecule type: mRNA  
A:Residues: 1-711 <CHO>  
A:Cross-references: EMBL:U07643; NID:G467236; PIDN:AAB60324.1; PID:G467237  
R:Rey, M.W.; Woloshuk, S.L.; deBoer, H.A.; Pieper, F.R.  
Nucleic Acids Res. 18, 5288, 1990  
A:Title: Complete nucleotide sequence of human mammary gland lactoferrin.  
A:Reference number: S11228; MUID:90384839; PMID:2402455  
A:Accession: S11228  
A:Molecule type: mRNA  
A:Residues: 1-148, 'T', 150-422, 'C', 424-711 <REY>  
A:Cross-references: EMBL:X53961; NID:G34415; PIDN:CAA37914.1; PID:G34416  
R:Teng, C.T.; Liu, Y.; Yang, N.; Walmer, D.; Panella, T.  
Mol. Endocrinol. 6, 1969-1981, 1992  
A:Title: Differential molecular mechanism of the estrogen action that regulates lactoferi  
A:Reference number: A45401; MUID:93125571; PMID:1480183  
A:Accession: A45401  
A:Molecule type: DNA  
A:Residues: 1-15 <TEN>  
A:Cross-references: GB:S52659; NID:G263311; PIDN:AAB24677.1; PID:G263312  
A:Experimental source: placenta  
A:Note: sequence extracted from NCBI backbone (NCBIIP:122202)  
R:Powell, M.J.; Ogden, J.E.  
Nucleic Acids Res. 18, 4013, 1990  
A:Title: Nucleotide sequence of human lactoferrin cDNA.  
A:Reference number: S10324; MUID:90326549; PMID:2374734  
A:Accession: S10324  
A:Molecule type: mRNA  
A:Residues: 3-711 <POW>  
A:Cross-references: EMBL:X52941; NID:G34411; PIDN:CAA37116.1; PID:G34412  
R:Stowell, K.M.; Rado, T.A.; Funk, W.D.; Tweedie, J.W.  
Biochem. J. 276, 349-355, 1991  
A:Title: Expression of cloned human lactoferrin in baby-hamster kidney cells.  
A:Reference number: S15853; MUID:91264786; PMID:2049056  
A:Accession: S15853  
A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 20-31 <ST1>  
A:Accession: S20841  
A:Molecule type: protein  
A:Residues: 20-28, X', 30-31 <ST2>

Query Match 71.0%; Score 49; DB 2; Length 708;  
Best Local Similarity 63.6%; Pred. No. 0.79;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWKRAMRKV 11  
| | | | |  
DB 38 CYQWQRNRKRL 48

RESULT 3  
S52107  
lactoferrin - sheep (fragment)  
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C>Date: 14-Jul-1995 #sequence\_revision 05-Dec-1997 #text\_change 05-Dec-1997  
C:Accession: S52107  
R:Qian, Z.Y.; Jolles, P.; Migliore-Samouri, D.; Fiat, A.M.  
Biochim. Biophys. Acta 1243, 25-32, 1995  
A>Title: Isolation and characterization of sheep lactoferrin, an inhibitor of platelet a  
A:Reference number: S52107; MUID:95127729; PMID:7827104  
A:Accession: S52107  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-33 <QA>  
C:Superfamily: transferrin; transferrin repeat homology  
C:Keywords: duplication

Query Match 66.7%; Score 46; DB 2; Length 33;  
Best Local Similarity 54.5%; Pred. No. 0.15; 1; Indels 0; Gaps 0;  
Matches 6; Conservative 4; Mismatches 0;

QY 1 CFQWKRAMRKV 11  
| | | | |  
DB 19 CYQWQRNRKRL 29

RESULT 4  
T22597  
hypothetical protein F53H4.4 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
C:Accession: T22597  
R:Dobson, R.  
submitted to the EMBL Data Library, October 1996  
A:Reference number: Z19587  
A:Accession: T22597  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-275 <WIL>  
A:Cross-references: EMBL:Z81089; PIDN:CAB03137.1; GSPDB:GN00028; CESP:F53H4.4  
A:Experimental source: clone F53H4  
C:Genetics:  
A:Gene: CESP:F53H4.4  
A:Map position: X  
A:Introns: 67/1; 153/1  
C:Superfamily: Caenorhabditis elegans hypothetical protein F53H4.4

Query Match 65.2%; Score 45; DB 2; Length 275;  
Best Local Similarity 72.7%; Pred. No. 1.7;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 FQWKRAMRKVR 12  
| | | | |  
DB 262 FQWKISWRKTR 272

RESULT 5  
H97451  
pyridoxamine 5'-phosphate oxidase (AF179611) [imported] - Agrobacterium tumefaciens (str.  
C:Species: Agrobacterium tumefaciens  
C>Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002  
C:Accession: H97451  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
C:Superfamily: H97451

Query Match 85.5%; Score 59; DB 1; Length 711;  
Best Local Similarity 83.3%; Pred. No. 0.013;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWKRAMRKVR 12  
| | | | |  
DB 39 CFQWQRNRKVR 50

RESULT 2  
JC2323  
lactoferrin - goat  
C:Species: Capra aegagrus hircus (domestic goat)  
C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 07-May-1999  
C:Accession: JC2323  
R:Le Provost, F.; Nocart, M.; Guerin, G.; Martin, P.  
Biochem. Biophys. Res. Commun. 203, 1324-1332, 1994  
A>Title: Characterization of the goat lactoferrin cDNA: assignment of the relevant locus  
A:Reference number: JC2323; MUID:94380047; PMID:8093048  
A:Accession: JC2323  
A:Molecule type: mRNA  
A:Residues: 1-708 <EP>  
C:Superfamily: transferrin; transferrin repeat homology  
C:Keywords: duplication; glycoprotein  
A:Status: predicted  
#status predicted

```

A;Experimental source: strain 16M
C;Genetics:
A;Gene: EMB11517
A;Map position: I
C;Superfamily: pyridoxamine-phosphate oxidase
C;Keywords: oxidoreductase

Query Match 62.3%; Score 43; DB 2; Length 208;
Best Local Similarity 58.3%; Pred. No. 2.9;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFQWKRAMRKVR 12
DB 90 CFHWKSLRQVR 101
|||:|:|
|||:|:|

RESULT 8
T14803
phytochrome C - sorghum
C;Species: Sorghum bicolor (sorghum)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 05-May-2000
C;Accession: T14803
R;Childs, K.B.; Miller, F.R.; Cordonnier-Pratt, M.M.; Pratt, L.H.; Morgan, P.W.; Mullet,
submitted to the EMBL Data Library, April 1996
A;Reference number: Z18186
A;Accession: T14803
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1135 <CHI>
A;Cross-references: EMBL:U56731; NID:G1800218; PID:G1800219
C;Genetics:
A;Gene: PHYC
C;Superfamily: phytochrome; phytochrome homology
C;Keywords: chromoprotein; dimer; photoreceptor; phytochromobilin; transcription regulati
F;65-581/Domain: phytochrome homology <PHY>
F;321/Binding site: phytochromobilin (Cys) (covalent) #status predicted

Query Match 60.9%; Score 42; DB 2; Length 1135;
Best Local Similarity 45.5%; Pred. No. 22;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWKRAMRKV 11
DB 775 CLEWNKAMQKI 785
||:|:|:|
||:|:|:|

RESULT 9
T20642
hypothetical protein F09C3.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T20642
R;Lennard, N.
submitted to the EMBL Data Library, March 1997
A;Reference number: Z19304
A;Accession: T20642
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1391 <WIL>
A;Cross-references: EMBL:Z92781; PIDN:CA807179.2; GSPDB:GN00019; CBSP:F09C3.1
A;Experimental source: clone F09C3
C;Genetics:
A;Gene: CESP:F09C3.1
A;Map position: 1
A;Introns: 53/2; 134/3; 283/1; 662/1; 731/3; 787/1; 976/3; 1043/2; 1171/3; 1267/2; 1354/3;

Query Match 59.4%; Score 41; DB 2; Length 1391;
Best Local Similarity 70.0%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 QWKRAMRKVR 12
DB 122 QWRKAMESVR 131
|||:|:|
|||:|:|

```

```

hypothetical protein YDR125c - Yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YD9727.20C
C:Species: Saccharomyces cerevisiae
C>Date: 19-May-1999 #sequence_revision 01-Sep-1999 #text_change 24-Sep-1999
A:Accession: S52690
R:Murphy, L.; Shore, L.; Harris, D.
Submitted to the EMBL Data Library, March 1995
A:Reference number: S52671
A:Accession: S52690
A:Molecule type: DNA
A:Residues: 1-453 <MUR>
A:Cross-references: EMBL:Z48758; NID:g747879; PIDN:CAA88678.1; PID:g747899; MIPS:YDR125C
C:Genetics:
A:Gene: SGD:ECM18
A:Cross-references: SGD:S0002532; MIPS:YDR125c
A:Map position: 4R
C:Superfamily: hypothetical protein YDR125c

Query Match 56.5%; Score 39; DB 2; Length 453;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CFQWKRAMRKVR 12
|:|:|:|:|
Db 24 CYWQKRAFHNSR 35

RESULT 13
T08030
dynein beta heavy chain - Chlamydomonas reinhardtii
C:Species: Chlamydomonas reinhardtii
C>Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 02-Feb-2001
C:Accession: T08030
R:Mitchell, D.R.; Brown, K.S.
J. Cell Sci. 107, 635-644, 1994
A:Title: Sequence analysis of the Chlamydomonas alpha and beta dynein heavy chain genes.
A:Reference number: Z16302; MUID:94274778; PMID:8006077
A:Accession: T08030
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-4568 <MIT>
A:Cross-references: EMBL:U02863; NID:g409965; PIDN:AAA19956.1; PID:g514215
A:Experimental source: strain 2igr
C:Genetics:
A:Gene: ODA4
A:Map position: IX
A:Introns: 48/2; 74/3; 105/3; 159/1; 225/3; 358/3; 514/3; 729/2; 919/3; 1004/3; 1115/3;
3334/3; 3686/3; 3882/3; 4240/3
C:Superfamily: dynein heavy chain, ciliary
C:Keywords: nucleotide binding; P-loop
F:1919-1926/Region: nucleotide-binding motif A (P-loop)
F:2202-2209/Region: nucleotide-binding motif A (P-loop)
F:2530-2537/Region: nucleotide-binding motif A (P-loop)

Query Match 56.5%; Score 39; DB 2; Length 4568;
Best Local Similarity 41.7%; Pred. No. 2.8e+02;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWKRAMRKVR 12
|:|:|:|:|
Db 1852 CFQWQSLRYIQ 1863

RESULT 14
A36415
32K protein - vaccinia virus (strain WR) (fragment)
C:Species: vaccinia virus
C>Date: 26-Jul-1991 #sequence_revision 26-Jul-1991 #text_change 21-Jul-2000
C:Accession: A36415
R:Pacha, R.F.; Weis, R.J.; Condit, R.C.
J. Virol. 64, 3853-3863, 1990
A:Title: Structure and expression of the vaccinia virus gene which prevents virus-induced
A:Reference number: A36415; MUID:90317884; PMID:2370683

```

---

```

roten F09C3.1 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
A:Accession: E87998
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biological
A:Reference number: A75000; MUID:99069613; PMID:9851916
Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: E87998
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1397 <STO>
A:Cross-references: GB:chr_I; PIDN:CAB07179.1; PID:g3875639; GSPDB:GN00019; CESP:F09C3.1
C:Genetics:
A:Gene: F09C3.1
A:Map position: 1

Query Match 59.4%; Score 41; DB 2; Length 1397;
Best Local Similarity 70.0%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Y 3 QWKRAMRKVR 12
|:|:|:|:|
b 122 QWRAMESVR 131

RESULT 11
28438
actoferrin precursor - mouse
N:Alternate names: lactotransferrin
C:Species: Mus musculus (house mouse)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
A:Accession: A28438; A41205
R:Pentecost, B.T.; Teng, C.T.
J. Biol. Chem. 262, 10134-10139, 1987
A:Title: Lactotransferrin is the major estrogen inducible protein of mouse uterine secretions
A:Reference number: A92596; MUID:87280033; PMID:3611056
A:Accession: A28438
A:Molecule type: mRNA
A:Residues: 3-707 <PEN>
A:Cross-references: EMBL:J03298
R:Liu, Y.; Teng, C.T.
J. Biol. Chem. 266, 21880-21885, 1991
A:Title: Characterization of estrogen-responsive mouse lactoferrin promoter.
A:Reference number: A41205; MUID:92042099; PMID:1939212
A:Accession: A41205
A:Molecule type: DNA
A:Residues: 1-15 <LIU>
A:Cross-references: GB:M74778
C:Superfamily: transferrin; transferrin repeat homology
C:Keywords: duplication; glycoprotein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-707/Product: lactotransferrin #status predicted <MAT>
F:358-695/Domain: transferrin repeat homology <RRH2>
F:494/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 58.0%; Score 40; DB 1; Length 707;
Best Local Similarity 54.5%; Pred. No. 32;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Y 1 CFQWKRAMRKV 11
|:|:|:|:|
b 37 CLRQWQNRKV 47

RESULT 12
152690

```



Accession: A36415  
Status: preliminary  
Molecule type: DNA  
Residues: 1-275 <PAC>  
Cross-references: EMBL:M32064; NID:G335834; PIDN:AAA48348.2; PID:G7555635

Query Match 55.1%; Score 38; DB 2; Length 275;  
Best Local Similarity 60.0%; Pred. No. 29;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

1 CFQWKRAMRK 10  
| : | | | |  
185 CLEWLRAXKK 194

RESULT 15  
8558  
Hypothetical protein Al7L - variola major virus  
Species: variola major virus  
Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 21-Jul-2000  
Accession: T28558  
Masung, R.F.; Esposito, J.J.; Liu, L.I.; Qi, J.; Utterback, T.R.; Knight, J.C.; Aubin  
ture 366, 748-751, 1993  
Title: Potential virulence determinants in terminal regions of variola smallpox virus  
Reference number: Z20488; MUID:94088747; PMID:8264798  
Accession: T28558  
Status: preliminary; translated from GB/EMBL/DBJ  
Molecule type: DNA  
Residues: 1-377 <MAS>  
Cross-references: EMBL:L22579; NID:G623595; PIDN:AAA60868.1; PID:G439038  
Experimental source: strain Bangladesh-1975

Query Match 55.1%; Score 38; DB 2; Length 377;  
Best Local Similarity 60.0%; Pred. No. 40;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

1 CFQWKRAMRK 10  
| : | | | |  
185 CLEWLRAXKK 194

Search completed: February 21, 2003, 08:02:48  
Search time : 12.6047 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

protein - protein search, using sw model

on: February 21, 2003, 07:28:06 ; Search time 5.2093 Seconds  
(without alignments) updates/sec  
95.544 Million cell

file: US-09-743-107B-93

rfect score: 69

quence: 1 CFQWKEMRKVR 12

oring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

arched: 112892 seqs, 41476328 residues

tal number of hits satisfying chosen parameters: 112892

imum DB seq length: 0

ximum DB seq length: 2000000000

st-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

atabase : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

sult	No.	Score	Query Match	Length	ID	Description
1	59	85.5	711	1	TRFL HUMAN	P02788 homo sapien
2	49	71.0	708	1	TRFL CAMDR	Q9tumo camelus dro
3	49	71.0	708	1	TRFL CAPHI	Q29477 capra hircu
4	42	60.9	1135	1	PHYC SORBI	P93528 sorghum bic
5	40	58.0	707	1	TRFL MOUSE	P08071 mus musculu
6	40	58.0	1137	1	PHYC ORISA	Q9zw19 oryza sativ
7	39	56.5	232	1	NLA DROME	Q9xz18 drosophila
8	39	56.5	4568	1	DYHE CHLRE	Q39565 chlamydomon
9	38	55.1	146	1	RPOB LIBAF	P41187 liberibacte
10	38	55.1	275	1	VA16 VACCV	P16710 vaccinia vi
11	38	55.1	378	1	VA16 VACCC	P20993 vaccinia vi
12	38	55.1	500	1	TLCE RICPR	C05962 rickettsia
13	38	55.1	856	1	ENV HV2NZ	P05883 human immun
14	38	55.1	2594	1	7LES DROVI	P20806 drosophila
15	37	53.6	528	1	CAX2 ARATH	Q38798 arabidopsis
16	37	53.6	530	1	CAX1 ARATH	P29402 arabidopsis
17	37	53.6	1131	1	PHY FINSY	Q41046 pinus sylve
18	36	52.2	133	1	D5 DICDI	P10382 dictyosteli
19	36	52.2	224	1	Y35A MYCGE	Q9zb72 mycoplasma
20	36	52.2	303	1	UL24 HSVSA	Q01005 herpesvirus
21	36	52.2	318	1	FX24 HUMAN	Q75426 homo sapien
22	36	52.2	502	1	C932 SOYBN	Q42799 glycine max
23	36	52.2	509	1	C931 SOYBN	Q42798 glycine max
24	36	52.2	510	1	C933 SOYBN	Q81973 glycine max
25	36	52.2	538	1	RO60 HUMAN	P10155 homo sapien
26	36	52.2	538	1	RO60 MOUSE	C00848 mus musculu
27	36	52.2	538	1	RO60 XENLA	P42700 xenopus lae
28	36	52.2	712	1	ENV HV2S2	P32536 human immun
29	36	52.2	859	1	ENV HV2ST	P20872 human immun
30	36	52.2	1111	1	PHYC ARATH	P14714 arabidopsis
31	36	52.2	1156	1	PHYE SOYBN	P42499 glycine max
32	36	52.2	1164	1	PHYD ARATH	P42497 arabidopsis
33	36	52.2	1172	1	PHYE ARATH	P14713 arabidopsis

34 36 52.2 2554 1 7LES DROME P13368 drosophila  
35 35 50.7 151 1 SYB2 RHIME Q923G1 rhizobium m  
36 35 50.7 264 1 CHL CANAL P33700 candida alb  
37 35 50.7 351 1 NOV CHICK P28686 gallus gall  
38 35 50.7 485 1 GLGA BACST O08328 bacillus st  
39 35 50.7 558 1 NCAP LYCVA P09992 lymphocytic  
40 35 50.7 558 1 NCAP LYCVA P07400 lymphocytic  
41 35 50.7 857 1 ENV HV2KR Q74126 human immun  
42 35 50.7 1121 1 PHY2 CERPU Q39557 ceratodon p  
43 35 50.7 1132 1 PHY1 PHYPA P36505 physcomitre  
44 35 50.7 1179 1 ATX1 ARATH Q9102 arabidopsis  
45 35 50.7 1214 1 TSGA RAT Q63679 rattus norv

#### ALIGNMENTS

RESULT 1  
TRFL\_HUMAN STANDARD; PRT; 711 AA.  
ID AC P02788; Q16780; Q16785; Q16789; Q00756; Q9H1Z3; Q96KZ4;  
AC Q96KZ5;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Lactotransferrin precursor (Lactoferrin) [Contains: Lactoferrroxin A;  
DE Lactoferrroxin B; Lactoferrroxin C].  
GN LFP OR LF.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RX MEDLINE=90384839; PubMed=2402455;  
RA Rey M.W., Woloshuk S.I., de Boer H.A., Pieper F.R.;  
RT "Complete nucleotide sequence of human mammary gland lactoferrin.";  
RL Nucleic Acids Res. 18:5288-5288(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RA Cho Y.Y.;  
RT Thesis (1994), Genetic Engineering Research Institute / Taejon, Korea.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC Corneely O.N.;  
RT Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RA Liang Q., Jimenez-Flores R., Richardson T.;  
RT "Molecular cloning and sequence analysis of human lactoferrin.";  
RN Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Bone marrow;  
RA Wei X., Han J., Rado T.A.;  
RT "Human neutrophil lactoferrin coding and 5' flanking region DNA  
sequences.";  
RN Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RA Cheng H., Chen X., Huan L.;  
RT "cDNA cloning and sequence analysis of human lactoferrin.";  
RN Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Prostate;  
RA Strausberg R.;  
RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
RN [8]

SEQUENCE OF 3-711 FROM N.A.  
P TISSUE=Mammary gland;  
C MEDLINE=90326549; PubMed=2374734;  
I Powell M.J., Ogden J.E., Mazurier J., Schoentgen F.,  
A "Nucleotide sequence of human lactoferrin cDNA.";  
L Nucleic Acids Res. 18:4013-4013(1990).  
[9]  
P SEQUENCE OF 20-711.  
X MEDLINE=85076667; PubMed=6510420;  
K Metz-Boutigue M.-H., Jolles J., Mazurier J., Schoentgen F.,  
A Legrand D., Spik G., Montreuil J., Jolles P.,  
I "Human lactotransferrin: amino acid sequence and structural  
L comparisons with other transferrins.";  
P Eur. J. Biochem. 145:659-666(1984).  
[10]  
P PRELIMINARY SEQUENCE OF 20-73; 134-171; 257-278 AND 347-530.  
X MEDLINE=82046817; PubMed=6794640;  
K Metz-Boutigue M.-H., Mazurier J., Jolles J., Spik G., Montreuil J.,  
A Jolles P.;  
I "The present state of the human lactotransferrin sequence. Study and  
L alignment of the cyanogen bromide fragments and characterization of  
P N- and C-terminal domains.";  
X Biochim. Biophys. Acta 670:243-254(1981).  
[11]  
P SEQUENCE OF 609-711.  
X MEDLINE=82262043; PubMed=7049727;  
K Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,  
A Jolles P.;  
I "An 88 amino acid long C-terminal sequence of human  
L lactotransferrin";  
P FEBS Lett. 142:107-110(1982).  
[12]  
P SEQUENCE OF 436-711 FROM N.A.  
X MEDLINE=88001031; PubMed=3477300;  
K Rado T.A., Wei X., Benz E.J. Jr.;  
I "Isolation of lactoferrin cDNA from a human myeloid library and  
L expression of mRNA during normal and leukemic myelopoiesis.";  
P Blood 70:989-993(1987).  
[13]  
P SEQUENCE OF 237-711 FROM N.A.  
X McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,  
A Nham M., Parrell L., Dedhia N., Ansari A., Mardis E., Schutz K.,  
A Gnoj L., la Bastide M., Kaplan N., Greco T., Touchman J., Muzny D.,  
C Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M., Porcel B.M.,  
A Dragan Y., Giacalone J., Pae A., Powell E., Solinsky K.A., Desilva U.,  
A Diaz-Perez S., Zhou X., Yu Y., Watanabe M., Doggett N., Garcia D.,  
L Segripanti J.L.;  
P Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
[14]  
P X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND REVISIONS.  
X MEDLINE=90064528; PubMed=2505506;  
A Anderson B.F., Baker H.M., Norris G.E., Rice D.W., Baker E.N.;  
I "Structure of human lactoferrin: crystallographic structure analysis  
L and refinement at 2.8-A resolution.";  
P J. Mol. Biol. 209:711-734(1989).  
[15]  
P X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
X Haridas M., Anderson B.F., Baker E.N.;  
I "Structure of human diferric lactoferrin refined at 2.2-A  
L resolution.";  
P Acta Crystallogr. D 51:629-646(1995).  
[16]  
P X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 20-353.  
X MEDLINE=97156796; PubMed=9003186;  
A Nicholson H., Anderson B.F., Bland T., Shewry S.C., Tweedie J.W.,  
A Baker E.N.;  
I "Maturation of the histidine ligand in human lactoferrin: iron  
L binding properties and crystal structure of the histidine-  
P 253->-methionine mutant.";  
X Biochemistry 36:341-346(1997).  
[17]  
P X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
X MEDLINE=99190892; PubMed=10089347;  
Sun X.L., Baker H.M., Shewry S.C., Jameson G.B., Baker E.N.;  
I "Structure of recombinant human lactoferrin expressed in Aspergillus  
L awamori.";  
P Acta Crystallogr. D 55:403-407(1999).  
[18]  
P X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
X MEDLINE=99192677; PubMed=10089508;  
K Jameson G.B., Anderson B.F., Norris G.E., Thomas D.H., Baker E.N.;  
I "Structure of human apolactoferrin at 2.0-A resolution. Refinement  
L and analysis of ligand-induced conformational change.";  
P Acta Crystallogr. D 54:1319-1335(1998).  
[19]  
P CHARACTERIZATION OF LACTOFERROXINS.  
X MEDLINE=91166929; PubMed=1369293;  
K Tani F., Iio K., Chiba H., Yoshikawa M.;  
I "Isolation and characterization of opioic antagonist peptides derived  
L from human lactoferrin.";  
P Agric. Biol. Chem. 54:1803-1810(1990).  
[20]  
P VARIANTS THR-30 AND ARG-48.  
X PubMed=9873069;  
K Klintworth G.K., Sommer J.R., Obrian G., Han L., Ahmed M.N.,  
A Qumsiyeh M.B., Lin P.-Y., Basti S., Reddy M.K., Kanai A., Hotta Y.,  
R Sugar J., Kumaramanickavel G., Munier F., Schorderet D.F.,  
A El Matri L., Iwata F., Kaiser-Kupfer M., Nagata M., Nakayasu K.,  
R Hejtmancik J.F., Teng C.T.;  
I "Familial subepithelial corneal amyloidosis (gelatinous drop-like  
L corneal dystrophy): exclusion of linkage to lactoferrin gene.";  
P Mol. Vision 4:31-32(1998).  
-!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
CC OF AN ANION, USUALLY BICARBONATE.  
-!- FUNCTION: LACTOFERROXINS A, B AND C HAVE OPIOID ANTAGONIST  
CC ACTIVITY. LACTOFERROXIN A SHOWS PREFERENCE FOR MU-RECEPTORS. WHILE  
CC LACTOFERROXIN B AND C HAVE SOMEWHAT HIGHER DEGREES OF PREFERENCE  
CC FOR KAPPA-RECEPTORS THAN FOR MU-RECEPTORS.  
-!- SUBUNIT: MONOMER.  
-!- SUBCELLULAR LOCATION: Secreted.  
-!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
-!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; X53961; CAA37914.1; -  
DR EMBL; U07643; AAB60324.1; -  
DR EMBL; M93150; AAA36159.1; -  
DR EMBL; M83202; AAA59511.1; -  
DR EMBL; M83205; AAA58656.1; -  
DR EMBL; M18642; AAA86665.1; -  
DR EMBL; AF332168; AAG48753.1; -  
DR EMBL; BC015822; AAH15822.1; -  
DR EMBL; BC015823; AAH15823.1; -  
DR EMBL; M73700; AAA59479.1; -  
DR EMBL; X52941; CAA37116.1; -  
DR EMBL; U95626; AAB57795.1; -  
DR PIR; S11228; TFHUL.  
DR PDB; 1LCF; 31-AUG-94.  
DR PDB; 1LCT; 31-OCT-93.  
DR PDB; 1LFG; 31-JUL-94.  
DR PDB; 1LFH; 31-OCT-93.  
DR PDB; 1LEI; 31-OCT-93.  
DR PDB; 1LGB; 31-AUG-94.  
DR PDB; 1LGC; 31-AUG-94.  
DR PDB; 1BXA; 08-NOV-96.  
DR PDB; 1DSN; 08-MAR-96.  
DR PDB; 1HSE; 12-MAR-97.  
DR PDB; 1VFD; 21-APR-97.

Query Match 85.5%; Score 59; DB 1; Length 711;  
 Best Local Similarity 83.3%; Pred. No. 0.0028;  
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 CFQWKRMRKVR 12  
 |||||  
 39 CFQWQRMRKVR 50

SULT 2  
 TRFL CAMDR STANDARD; PRT; 708 AA.  
 Q9TUM0; Q9WZS5;  
 16-OCT-2001 (Rel. 40, Created)  
 16-OCT-2001 (Rel. 40, Last sequence update)  
 15-JUN-2002 (Rel. 41, Last annotation update)  
 Lactotransferrin precursor (Lactoferrin).  
 LTF.  
 Camelus dromedarius (Dromedary) (Arabian camel).  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.  
 NCBI\_TaxID=9838;  
 [1]  
 SEQUENCE FROM N.A.  
 STRAIN=Sonali; TISSUE=lactating mammary gland;  
 Kappeler S.R.; Ackermann M.; Farah Z.; Fuhan Z.;  
 "Sequence analysis of camel (Camelus dromedarius) lactoferrin."  
 Int. Dairy J. 9:481-486(1999).  
 [2]  
 SEQUENCE FROM N.A.  
 TISSUE=Mammary gland;  
 Paramasivam M.; Srinivasan A.; Singh R.; Sahani M.S.; Singh T.P.;  
 Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 OF AN ANION, USUALLY BICARBONATE (BY SIMILARITY).  
 -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 -!- SUBCELLULAR LOCATION: Secreted.  
 -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 -----

EMBL; AJ131674; CAB53387.1; -;  
 EMBL; AF165879; AAF82241.1; -;  
 HSSP; Q77811; 1B1X.  
 InterPro; IPR001156; Transferrin.  
 Pfam; PF00405; transferrin; 2.  
 PRINTS; PR00422; TRANSFERRIN.  
 SMART; SM00094; TR\_FER; 2.  
 PROSITE; PS00205; TRANSFERRIN\_1; 2.  
 PROSITE; PS00206; TRANSFERRIN\_2; 2.  
 PROSITE; PS00207; TRANSFERRIN\_3; 2.  
 Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 Signal.  
 SIGNAL  
 CHAIN 1 19 BY SIMILARITY.  
 REPEAT 20 708 LACTOTRANSFERRIN.  
 REPEAT 20 363 1.  
 REPEAT 364 708 2.  
 DISULFID 28 64 BY SIMILARITY.  
 DISULFID 38 55 BY SIMILARITY.  
 DISULFID 134 217 BY SIMILARITY.  
 DISULFID 176 192 BY SIMILARITY.  
 DISULFID 189 200 BY SIMILARITY.  
 DISULFID 250 264 BY SIMILARITY.  
 DISULFID 367 399 BY SIMILARITY.

FT DISULFID 377 390 BY SIMILARITY.  
 FT DISULFID 424 703 BY SIMILARITY.  
 FT DISULFID 444 666 BY SIMILARITY.  
 FT DISULFID 476 551 BY SIMILARITY.  
 FT DISULFID 500 694 BY SIMILARITY.  
 FT DISULFID 510 524 BY SIMILARITY.  
 FT DISULFID 521 534 BY SIMILARITY.  
 FT DISULFID 592 606 BY SIMILARITY.  
 FT DISULFID 644 649 BY SIMILARITY.  
 FT METAL 79 79 IRON 1 (BY SIMILARITY).  
 FT METAL 111 111 IRON 1 (BY SIMILARITY).  
 FT METAL 211 211 IRON 1 (BY SIMILARITY).  
 FT METAL 272 272 IRON 1 (BY SIMILARITY).  
 FT METAL 414 414 IRON 2 (BY SIMILARITY).  
 FT METAL 452 452 IRON 2 (BY SIMILARITY).  
 FT METAL 545 545 IRON 2 (BY SIMILARITY).  
 FT METAL 614 614 IRON 2 (BY SIMILARITY).  
 FT BINDING 140 140 ANION (BY SIMILARITY).  
 FT BINDING 482 482 ANION (BY SIMILARITY).  
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 385 385 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 537 537 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 594 594 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CONFLICT 261 261 F -> S (IN REF. 2).  
 FT CONFLICT 304 304 G -> A (IN REF. 2).  
 FT CONFLICT 330 330 S -> P (IN REF. 2).  
 FT CONFLICT 492 494 LLS -> PLF (IN REF. 2).  
 FT CONFLICT 506 506 L -> F (IN REF. 2).  
 FT CONFLICT 609 609 A -> P (IN REF. 2).  
 FT CONFLICT 642 642 R -> Q (IN REF. 2).  
 SQ SEQUENCE 708 AA; 77211 MW; 0B0C175A0B69D430 CRC64;

Query Match 71.0%; Score 49; DB 1; Length 708;  
 Best Local Similarity 66.7%; Pred. No. 0.19;  
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWKRMRKVR 12  
 |||||  
 DB 38 CAQWQRMRKVR 49

RESULT 3  
 TRFL CAPHI STANDARD; PRT; 708 AA.  
 ID TRFL CAPHI STANDARD; PRT; 708 AA.  
 AC Q29477; Q29479;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin).  
 GN LTF.  
 OS Capra hircus (Goat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Capra.  
 OX NCBI\_TaxID=9925;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Lee T.; Yu S.; Kim S.; Lee K.; Yu D.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RX MEDLINE=94380047; PubMed=8093048;  
 RA le Provost F.; Nocard M.; Guerin G.; Martin P.;  
 RT "Characterization of the goat lactoferrin cDNA. Assignment of the  
 relevant locus to bovine U12 syntenic group."  
 RL Biochem Biophys Res Commun. 203:1324-1332(1994).  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 OF AN ANION, USUALLY BICARBONATE.  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Secreted.

C -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 C -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
 C  
 C This SWISS-PROT entry is copyright. It is produced through a collaboration  
 C between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 C the European Bioinformatics Institute. There are no restrictions on its  
 C use by non-profit institutions as long as its content is in no way  
 C modified and this statement is not removed. Usage by and for commercial  
 C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 C or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 C  
 C -----  
 C EMBL; U53857; AAA97958.1; --  
 C EMBL; X78902; CAAS5517.1; --  
 C HSP; O77698; ICE2.  
 C InterPro; IPR001156; Transferrin.  
 C Pfam; PF00405; transferrin; 2.  
 C PRINTS; PR00422; TRANSFERRIN.  
 C SMART; SM00094; TF\_FER; 2.  
 C PROSITE; PS00205; TRANSFERRIN 1; 2.  
 C PROSITE; PS00206; TRANSFERRIN 2; 2.  
 C PROSITE; PS00207; TRANSFERRIN 3; 2.  
 C Transprot; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 C Signal.  
 C  
 C 1 19 BY SIMILARITY  
 C CHAIN 20 708 LACTOTRANSFERRIN.  
 C REPEAT 20 363 1.  
 C REPEAT 364 708 2.  
 C DISULFID 28 64 BY SIMILARITY.  
 C DISULFID 38 55 BY SIMILARITY.  
 C DISULFID 134 217 BY SIMILARITY.  
 C DISULFID 176 192 BY SIMILARITY.  
 C DISULFID 189 200 BY SIMILARITY.  
 C DISULFID 250 264 BY SIMILARITY.  
 C DISULFID 367 399 BY SIMILARITY.  
 C DISULFID 377 390 BY SIMILARITY.  
 C DISULFID 424 703 BY SIMILARITY.  
 C DISULFID 444 665 BY SIMILARITY.  
 C DISULFID 476 551 BY SIMILARITY.  
 C DISULFID 500 694 BY SIMILARITY.  
 C DISULFID 510 524 BY SIMILARITY.  
 C DISULFID 521 534 BY SIMILARITY.  
 C DISULFID 592 606 BY SIMILARITY.  
 C DISULFID 644 649 BY SIMILARITY.  
 C METAL 79 79 IRON 1 (BY SIMILARITY).  
 C METAL 111 111 IRON 1 (BY SIMILARITY).  
 C METAL 211 211 IRON 1 (BY SIMILARITY).  
 C METAL 272 272 IRON 1 (BY SIMILARITY).  
 C METAL 414 414 IRON 2 (BY SIMILARITY).  
 C METAL 452 452 IRON 2 (BY SIMILARITY).  
 C METAL 545 545 IRON 2 (BY SIMILARITY).  
 C METAL 614 614 IRON 2 (BY SIMILARITY).  
 C BINDING 140 140 ANION (BY SIMILARITY).  
 C BINDING 482 482 ANION (BY SIMILARITY).  
 C CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 C CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 C CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 C CARBOHYD 495 495 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 C CARBOHYD 564 564 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 C CONFLICT 56 56 I -> V (IN REF. 2).  
 C CONFLICT 88 88 L -> R (IN REF. 2).  
 C CONFLICT 124 124 Q -> K (IN REF. 2).  
 C CONFLICT 154 154 P -> P (IN REF. 2).  
 C CONFLICT 304 304 S -> R (IN REF. 2).  
 C CONFLICT 414 414 D -> G (IN REF. 2).  
 C SEQUENCE 708 AA; 77358 MW; F2EDA3CB3539960D CRC64;  
 C  
 C Query Match 71.08; Score 49; DB 1; Length 708;  
 C Best Local Similarity 63.6%; Pred. No. 0.19;  
 C Matches 7; Conservative 3; Mismatches 0; Gaps 0;  
 C  
 C 1 CFQWKRAMRKV 11  
 C : : : : :  
 C 38 CYQWRMRKL 48

## RESULT 4

PHYC\_SORBI STANDARD; PRT; 1135 AA.  
 ID PHYC\_SORBI  
 AC P93528;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Phytochrome C.  
 GN PHYC.  
 OS Sorghum bicolor (Sorghum) (Sorghum vulgare).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; PACC Clade;  
 OC Panicoideae; Andropogoneae; Sorghum.  
 OX NCBI\_TaxID=4558;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97198556; PubMed=9046599;  
 RA Childs K.L., Miller F.R., Cordonnier-Pratt M.-M., Pratt L.H.,  
 RA Morgan P.W., Mullet J.E.;  
 RT "The Sorghum bicolor photoperiod sensitivity gene, Ma3, encodes a  
 RT phytochrome B.";  
 RL Plant Physiol. 113:611-619(1997).  
 CC -1- FUNCTION: REGULATORY PHOTORECEPTOR WHICH EXISTS IN TWO FORMS THAT  
 CC ARE REVERSIBLY INTERCONVERTIBLE BY LIGHT: THE PR FORM THAT ABSORBS  
 CC MAXIMALLY IN THE RED REGION OF THE SPECTRUM AND THE PFR FORM THAT  
 CC ABSORBS MAXIMALLY IN THE FAR-RED REGION. PHOTOCONVERSION OF PFR IN  
 CC PFR INDUCES AN ARRAY OF MORPHOGENIC RESPONSES, WHEREAS  
 CC RECONVERSION OF PFR TO PR CANCELS THE INDUCTION OF THOSE  
 CC RESPONSES. PFR CONTROLS THE EXPRESSION OF A NUMBER OF NUCLEAR  
 CC GENES INCLUDING THOSE ENCODING THE SMALL SUBUNIT OF RIBULOSE-  
 CC BISPHOSPHATE CARBOXYLASE, CHLOROPHYLL A/B BINDING PROTEIN,  
 CC PROCHLOROPHYLLIDE REDUCTASE, RNA, ETC. IT ALSO CONTROLS THE  
 CC EXPRESSION OF ITS OWN GENE(S) IN A NEGATIVE FEEDBACK FASHION (BY  
 CC SIMILARITY).  
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -1- PTM: CONTAINS ONE COVALENTLY LINKED TETRAPYRROLE CHROMOPHORE.  
 CC -1- SIMILARITY: BELONGS TO THE PHYTOCHROME FAMILY.  
 CC -1- SIMILARITY: CONTAINS 2 PAS (PER-ARNT-SIM) DIMERIZATION DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 CC EMBL; U56731; AAB41399.1; --  
 CC InterPro; IPR003594; ATPbind\_ATPase.  
 CC InterPro; IPR003018; GAF.  
 CC InterPro; IPR004359; HIS\_KIN\_sig.  
 CC InterPro; IPR003661; His\_KinA.  
 CC InterPro; IPR001610; PAC.  
 CC InterPro; IPR000014; PAS\_domain.  
 CC InterPro; IPR001294; Phytochrome.  
 CC Pfam; PF00360; phytochrome; 1.  
 CC Pfam; PF00512; signal; 1.  
 CC Pfam; PF00389; PAS; 2.  
 CC Pfam; PF01590; GAF; 1.  
 CC Pfam; PF02518; HATPase\_c; 1.  
 CC Pfam; PF01033; PHYTOCHROME.  
 CC SMART; SM00065; GAF; 1.  
 CC SMART; SM00387; HATPase\_c; 1.  
 CC SMART; SM00388; HisK\_A; 1.  
 CC SMART; SM00086; PAC; 1.  
 CC SMART; SM00091; PAS; 2.  
 CC TIGRFAMs; TIGR00229; sensory\_box; 2.  
 CC PROSITE; PS0109; HIS\_KIN; 1.  
 CC PROSITE; PS0112; PAS; 2.

```
PROSITE; PS00245; PHYTOCHROME_1; FALSE NEG.
PROSITE; PS00046; PHYTOCHROME_2; 1.
Transcription regulation; Photoreceptor; Phytochrome; Chromophore;
Repeat; Multigene family.
DOMAIN 618 688 PAS 1.
DOMAIN 748 822 PAS 2.
DOMAIN 902 1122 HISTIDINE KINASE.
BINDING 321 321 CHROMOPHORE (BY SIMILARITY).
SEQUENCE 1135 AA; 126315 NW; AFCC934B7592DE4D CRC64;

Query Match 60.9%; Score 42; DB 1; Length 1135;
Best Local Similarity 45.5%; Pred. No. 6.2;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

1 CFQWKRMRKV 11
| : : : :
775 CLEWKMQRKI 785

SULT 5
FL_MOUSE STANDARD; PRT; 707 AA.
P08071; P70690; Q61799; Q92282;
01-AUG-1998 (Rel. 08, Created)
15-JUN-2002 (Rel. 41, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
Lactoferrin precursor (Lactoferrin).
LTF.
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
TSSUE=Uterus;
MEDLINE=87280033; PubMed=3611056;
Pentecost B.T.; Teng C.T.;
"Lactoferrin is the major estrogen inducible protein of mouse
uterine secretions.";
J. Biol. Chem. 262:10134-10139(1987).
[2]
SEQUENCE FROM N.A.
TSSUE=Uterus;
Morihashi K.;
Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
Straussberg R.;
Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE OF 1-14 FROM N.A.
MEDLINE=92042099; PubMed=1939212;
Liu Y.; Teng C.T.;
"Characterization of estrogen-responsive mouse lactoferrin promoter.";
J. Biol. Chem. 266:21880-21885(1991).
-!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
OF AN ANION, USUALLY BICARBONATE.
-!- SUBUNIT: MONOMER.
-!- SUBCELLULAR LOCATION: Secreted.
-!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
-!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; J03298; AAA40525.1; -.
EMBL; D88510; BAA13633.1; -.

DR EMBL; BC006904; AAA06904.1; -.
DR EMBL; M74778; AAA39427.1; -.
DR PIR; A28438; A28438.
DR HSP; P02788; 1CB6.
DR MGD; MGI:96837; Ltf.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 2.
DR PRINTS; PR00422; TRANSFERRIN.
DR SMART; SMO0094; TR_FER; 2.
DR PROSITE; PS00205; TRANSFERRIN_1; 1.
DR PROSITE; PS00206; TRANSFERRIN_2; 2.
DR PROSITE; PS00207; TRANSFERRIN_3; 2.
KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;
KW Signal.
FT CHAIN 1 19 BY SIMILARITY.
FT REPEAT 20 707 LACTOTRANSFERRIN.
FT REPEAT 20 357 1.
FT REPEAT 358 707 2.
FT DISULFID 27 63 BY SIMILARITY.
FT DISULFID 37 54 BY SIMILARITY.
FT DISULFID 133 216 BY SIMILARITY.
FT DISULFID 175 191 BY SIMILARITY.
FT DISULFID 188 199 BY SIMILARITY.
FT DISULFID 249 263 BY SIMILARITY.
FT DISULFID 366 398 BY SIMILARITY.
FT DISULFID 376 389 BY SIMILARITY.
FT DISULFID 702 707 BY SIMILARITY.
FT DISULFID 423 443 BY SIMILARITY.
FT DISULFID 443 665 BY SIMILARITY.
FT DISULFID 475 550 BY SIMILARITY.
FT DISULFID 499 693 BY SIMILARITY.
FT DISULFID 509 523 BY SIMILARITY.
FT DISULFID 520 533 BY SIMILARITY.
FT DISULFID 591 605 BY SIMILARITY.
FT DISULFID 643 648 BY SIMILARITY.
FT METAL 78 78 IRON 1 (BY SIMILARITY).
FT METAL 110 110 IRON 1 (BY SIMILARITY).
FT METAL 210 210 IRON 1 (BY SIMILARITY).
FT METAL 271 271 IRON 1 (BY SIMILARITY).
FT METAL 413 413 IRON 2 (BY SIMILARITY).
FT METAL 451 451 IRON 2 (BY SIMILARITY).
FT METAL 544 544 IRON 2 (BY SIMILARITY).
FT METAL 613 613 IRON 2 (BY SIMILARITY).
FT BINDING 139 139 ANION (POTENTIAL).
FT BINDING 481 481 ANION (POTENTIAL).
FT CARBOHYD 118 118 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 494 494 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 1 2 MR -> IQG (IN REF. 1).
FT CONFLICT 25 25 R -> Q (IN REF. 2).
FT CONFLICT 82 82 M -> L (IN REF. 2).
FT CONFLICT 359 359 S -> T (IN REF. 2).
FT CONFLICT 382 382 A -> D (IN REF. 1).
FT CONFLICT 449 449 E -> G (IN REF. 2).
FT CONFLICT 629 629 L -> V (IN REF. 1).
SQ SEQUENCE 707 AA; 77865 MW; F26AE0340AAC19A8 CRC64;

Query Match 58.0%; Score 40; DB 1; Length 707;
Best Local Similarity 54.5%; Pred. No. 8.8;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWKRMRKV 11
| : : : :
DB 37 CLRQWQEMRKV 47

RESULT 6
PHYC ORYSA
ID PHYC ORYSA STANDARD; PRT; 1137 AA.
AC Q9ZW19; P93429;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phytochrome C.
GN PHYC.
```

S Oryza sativa (Rice).  
 S Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 S Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 S Ehrhartoideae; Oryzaceae; Oryza.  
 S NCBI\_TaxID=4530;  
 S [1]  
 S SEQUENCE FROM N.A.  
 S STRAIN=cv. Nipponbare;  
 S Tahir M., Kangas H., Takano M.;  
 S "Phytochrome C (PHYC) gene in rice: isolation and characterization of  
 S a complete coding sequence.";  
 S (in) Plant Gene Register PGR98-210.  
 S [2]  
 S SEQUENCE OF 275-378 FROM N.A.  
 S MEDLINE=97019052; PubMed=8965668;  
 S Matthews S., Sharrock R.A.;  
 S "The phytochrome gene family in grasses (Poaceae): a phylogeny and  
 S evidence that grasses have a subset of the loci found in dicot  
 S angiosperms.";  
 S Mol. Biol. Evol. 13:1141-1150(1996).  
 S -!- FUNCTION: REGULATORY PHOTORECEPTOR WHICH EXISTS IN TWO FORMS THAT  
 S ARE REVERSIBLY INTERCONVERTIBLE BY LIGHT: THE PR FORM THAT ABSORBS  
 S MAXIMALLY IN THE RED REGION OF THE SPECTRUM AND THE PFR FORM THAT  
 S ABSORBS MAXIMALLY IN THE FAR-RED REGION. PHOTOCONVERSION OF PFR IN  
 S PFR INDUCES AN ARRAY OF MORPHOGENIC RESPONSES, WHEREAS  
 S RECONVERSION OF PFR TO PR CANCELS THE INDUCTION OF THOSE  
 S RESPONSES. PFR CONTROLS THE EXPRESSION OF A NUMBER OF NUCLEAR  
 S GENES INCLUDING THOSE ENCODING THE SMALL SUBUNIT OF RUBULOSE-  
 S BIPHOSPHATE CARBOXYLASE, CHLOROPHYLL A/B BINDING PROTEIN,  
 S PROTOCHLOROPHYLLIDE REDUCTASE, RNA, ETC. IT ALSO CONTROLS  
 S THE EXPRESSION OF ITS OWN GENE(S) IN A NEGATIVE FEEDBACK FASHION.  
 S -!- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 S -!- PTM: CONTAINS ONE COVALENTLY LINKED TETRAPYRROLE CHROMOPHORE.  
 S -!- SIMILARITY: BELONGS TO THE PHYTOCHROME FAMILY.  
 S -!- SIMILARITY: CONTAINS 2 PAS (PER-ARNT-SIM) DIMERIZATION DOMAINS.  
 S -!- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.  
 S -!- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.  
 S -----  
 S This SWISS-PROT entry is copyright. It is produced through a collaboration  
 S between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 S the European Bioinformatics Institute. There are no restrictions on its  
 S use by non-profit institutions as long as its content is in no way  
 S modified and this statement is not removed. Usage by and for commercial  
 S entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 S or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 S -----  
 S EMBL; AB018442; BAA74448.1; -  
 S EMBL; U61207; AAB41996.1; -  
 S InterPro; IPR003594; ATPbind\_ATPase.  
 S InterPro; IPR003018; GAF.  
 S InterPro; IPR004359; HIS\_KIN\_sig.  
 S InterPro; IPR003661; His\_kinA.  
 S InterPro; IPR001610; PAC.  
 S InterPro; IPR000014; PAS domain.  
 S InterPro; IPR001294; Phytochrome.  
 S Pfam; PF00360; phytochrome; 1.  
 S Pfam; PF00512; signal; 1.  
 S Pfam; PF00389; PAS; 2.  
 S Pfam; PF00390; GAF; 1.  
 S Pfam; PF02518; HATPase\_c; 1.  
 S PRINTS; PF01033; PHYTOCHROME.  
 S SMART; SM00065; GAF; 1.  
 S SMART; SM00387; HATPase\_c; 1.  
 S SMART; SM00388; HisK; 1.  
 S SMART; SM00086; PAC; 1.  
 S SMART; SM00091; PAS; 2.  
 S TIGRFAMs; TIGR00229; sensory\_box; 2.  
 S PROSITE; PS50109; HIS\_KIN; 1.  
 S PROSITE; PS50112; PAS; 2.  
 S PROSITE; PS00245; PHYTOCHROME\_1; 1.  
 S PROSITE; PS50046; PHYTOCHROME\_2; 1.  
 S Transcription regulation; Photoreceptor; Phytochrome; Chromophore;  
 S Repeat; Multigene family.

FT DOMAIN 620 690 PAS 1.  
 FT DOMAIN 750 824 PAS 2.  
 FT BINDING 904 1124 HISTIDINE KINASE.  
 FT BINDING 322 322 CHROMOPHORE (BY SIMILARITY).  
 FT CONFLICT 279 279 F -> S (IN REF. 2).  
 FT CONFLICT 292 292 C -> S (IN REF. 2).  
 SQ SEQUENCE 1137 AA; 125982 MW; F2A520181CFE7B32 CRC64;  
 Query Match 58.0%; Score 40; DB 1; Length 1137;  
 Best Local Similarity 45.5%; Pred. No. 15;  
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 CFQWKRVRKV 11  
 DB 777 CLEWNEAQKI 787  
 RESULT 7  
 NLA\_DROME  
 ID NLA\_DROME STANDARD; PRT; 292 AA.  
 AC Q9XZL8; Q9V391;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Nebula protein.  
 GN NLA OR CG6072.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA McCormick A.V., Goldberg M.L.;  
 RT "Gene required for elongation of meiosis I spindle in Drosophila  
 RT females.";  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Calle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter B.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,  
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Fabros B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Flossler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodier A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jaisli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Turner C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,

Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 "The genome sequence of *Drosophila melanogaster*."  
 Science 287:2185-2195(2000).  
 -1- FUNCTION: REQUIRED FOR ELONGATION OF MEIOSIS I SPINDLE.  
 -1- SIMILARITY: BELONGS TO THE DSCR1 FAMILY.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; AF147700; AAD33987.1; -.  
 EMBL; AF003712; AAF5285.1; -.  
 Flybase; FBgn0026629; nla.  
 SEQUENCE 292 AA; 31423 MW; 64F1BBF5F6EA6CF9 CRC64;  
 -----  
 Query Match 56.5%; Score 39; DB 1; Length 292;  
 Best Local Similarity 54.5%; Pred. No. 5.3; Mismatches 2; Indels 0; Gaps 0;  
 Matches 6; Conservative 3;  
 -----  
 2 F0WKRMRKVR 12  
 ||| | : : :  
 150 F0WLRSPRLR 160  
 -----  
 SULT 8  
 HB CHLRE  
 DYHE CHLRE STANDARD; PRT; 4568 AA.  
 Q39565;  
 01-NOV-1997 (Rel. 35, Created)  
 01-NOV-1997 (Rel. 35, Last sequence update)  
 30-MAY-2000 (Rel. 39, Last annotation update)  
 Dynein beta chain, flagellar outer arm.  
 ODA4 OR ODA-4 OR SUP1.  
 Chlamydomonas reinhardtii.  
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
 Chlamydomonadaceae; Chlamydomonas.  
 NCBI\_TaxID=3055;  
 [1]  
 -----  
 SEQUENCE FROM N.A.  
 STRAIN=21gr;  
 MEDLINE=94274778; PubMed=8006077;  
 Mitchell D.R., Brown K.S.;  
 "Sequence analysis of the Chlamydomonas alpha and beta dynein heavy chain genes".  
 J. Cell Sci. 107:635-644(1994).  
 -1- FUNCTION: FORCE GENERATING PROTEIN OF EUKARYOTIC CILIA AND FLAGELLA. PRODUCES FORCE TOWARDS THE MINUS ENDS OF MICROTUBULES. DYNEIN HAS ATPASE ACTIVITY.  
 -1- SUBUNIT: CONSISTS OF AT LEAST 3 HEAVY CHAINS (ALPHA, BETA AND GAMMA), 2 INTERMEDIATE CHAINS AND 8 LIGHT CHAINS.  
 -1- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; U02963; AAA1956.1; -.  
 InterPro; IPR004273; Dynein heavy.  
 Pfam; PF03028; Dynein heavy; 1.  
 Motor protein; Microtubules; Dynein; ATP-binding; Flagella;  
 Coiled coil.  
 DOMAIN 277 293 COILED COIL (POTENTIAL).  
 -----

FT DOMAIN 1158 1175 COILED COIL (POTENTIAL).  
 FT DOMAIN 1372 1400 COILED COIL (POTENTIAL).  
 FT DOMAIN 1614 1630 COILED COIL (POTENTIAL).  
 FT DOMAIN 1778 1825 COILED COIL (POTENTIAL).  
 FT DOMAIN 2017 2045 MICROTUBULE-BINDING (POTENTIAL).  
 FT DOMAIN 2831 2848 COILED COIL (POTENTIAL).  
 FT DOMAIN 3106 3162 COILED COIL (POTENTIAL).  
 FT DOMAIN 3339 3425 COILED COIL (POTENTIAL).  
 FT DOMAIN 3648 3728 COILED COIL (POTENTIAL).  
 FT NP\_BIND 1919 1926 ATP (POTENTIAL).  
 FT NP\_BIND 2202 2209 ATP (POTENTIAL).  
 FT NP\_BIND 2530 2537 ATP (POTENTIAL).  
 FT NP\_BIND 2879 2886 ATP (POTENTIAL).  
 SQ SEQUENCE 4568 AA; 519961 MW; 9A9A5393C7C36AE7 CRC64;  
 -----  
 Query Match 56.5%; Score 39; DB 1; Length 4568;  
 Best Local Similarity 41.7%; Pred. No. 96;  
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
 -----  
 Qy 1 CFQWKRMRKVR 12  
 ||| | : : :  
 Db 1852 CFQWQSQLRYIQ 1863  
 -----  
 RESULT 9  
 RPOB LIBAF  
 ID RPOB LIBAF STANDARD; PRT; 146 AA.  
 AC P41187;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE DNA-directed RNA polymerase beta chain (EC 2.7.7.6) (Transcriptase beta chain) (RNA polymerase beta subunit) (Fragment).  
 GN RPOB.  
 OS Liberibacter africanus (Liberibacter africanum).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Candidatus Liberibacter.  
 OX NCBI\_TaxID=34020;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Nelspruit;  
 RA Planet P., Jagoueix S., Bove J.M., Garnier M.;  
 RT "Detection and characterization of the African Citrus Greening Liberibacter by amplification, cloning and sequencing of the rplKALJL-rpoBC operon".  
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS SUBSTRATES.  
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA)(N).  
 CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1 BETA' CHAIN.  
 CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; U09675; AAA19557.1; -.  
 InterPro; IPR001572; RNA\_pol\_B.  
 Pfam; PF00562; RNA\_pol\_B; 1.  
 DR PROSITE; PS01166; RNA\_POL\_BETA; PARTIAL.  
 KW Transferase; Transcription; DNA-directed RNA polymerase.  
 FT NON\_TER 146 146  
 SQ SEQUENCE 146 AA; 16731 MW; 815D6EDD66FD8943 CRC64;  
 -----  
 Query Match 55.1%; Score 38; DB 1; Length 146;



Best Local Similarity 60.0%; Pred. No. 3.9;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Y 1 CFQWKRAMRK 10  
| | | | |  
b 10 CQWMSGARK 19  
| | | | |

RESULT 10  
A16\_VACCC STANDARD; PRT; 275 AA.  
C P16710;  
T 01-FEB-1991 (Rel. 15, Created)  
T 01-AUG-1990 (Rel. 15, Last sequence update)  
T 16-OCT-2001 (Rel. 40, Last annotation update)  
T Protein A16 (Fragment).  
X A16L.  
S Vaccinia virus (strain WR).  
C Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
C Orthopoxvirus.  
C NCBI\_TaxID=10254;  
X [1]  
P SEQUENCE FROM N.A.  
K MEDLINE=90317684; PubMed=2370693;  
A Facha R.F., Meis R.J., Condit R.C.;  
T "Structure and expression of the vaccinia virus gene which prevents  
T virus-induced breakdown of RNA."  
J. Virol. 64:3853-3863(1990).  
L -1- SIMILARITY: BELONGS TO THE POXVIRUSES A16 FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; M32064; AAA48348.2; -  
PIR; A36415; A36415.  
R InterPro; IPR004251; DUF230.  
R Pfam; PF03003; DUF230; 1.  
T NON TER 275  
Q SEQUENCE 275 AA; 31811 MW; E2461AB1DB7B93A3 CRC64;  
-----  
Query Match 55.1%; Score 38; DB 1; Length 275;  
Best Local Similarity 60.0%; Pred. No. 7.6;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Y 1 CFQWKRAMRK 10  
| | | | |  
b 185 CLEWLAKRK 194  
| | | | |

RESULT 11  
A16\_VACCC STANDARD; PRT; 378 AA.  
C P20953;  
T 01-FEB-1991 (Rel. 17, Created)  
T 01-FEB-1991 (Rel. 17, Last sequence update)  
T 16-OCT-2001 (Rel. 40, Last annotation update)  
T Protein A16.  
X A16L.  
S Vaccinia virus (strain Copenhagen).  
C Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
C Orthopoxvirus.  
C NCBI\_TaxID=10249;  
X [1]  
P SEQUENCE FROM N.A.  
K MEDLINE=91021027; PubMed=2219722;  
A Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,  
A Paolletti E.;  
T "The complete DNA sequence of vaccinia virus."

Virology 179:247-266(1990).  
[2]  
RN COMPLETE GENOME.  
RP Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,  
RA Paolletti E.;  
RT "Appendix to 'The complete DNA sequence of vaccinia virus'";  
RL Virology 179:517-563(1990).  
CC -1- SIMILARITY: BELONGS TO THE POXVIRUSES A16 FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; M35027; AAA48138.1; -  
PIR; I42518; I42518.  
R InterPro; IPR004251; DUF230.  
R Pfam; PF03003; DUF230; 1.  
S SEQUENCE 378 AA; 43561 MW; 05ED614AA1D11A19 CRC64;  
-----  
Query Match 55.1%; Score 38; DB 1; Length 378;  
Best Local Similarity 60.0%; Pred. No. 11;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWKRAMRK 10  
| | | | |  
Db 185 CLEWLAKRK 194  
| | | | |

RESULT 12  
TLCE RICPR STANDARD; PRT; 500 AA.  
ID TLCE RICPR  
AC OC5962;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE ADP/ATP carrier protein 5 (ADP/ATP translocase 5).  
GN TLCE OR TLC5 OR RP739.  
OS Rickettsia prowazekii.  
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;  
OC Rickettsiaceae; Rickettsiae; Rickettsia.  
OX NCBI\_TaxID=782;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Madrid E;  
RX MEDLINE=99039499; PubMed=9823893;  
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,  
RA Sickeritz-Ponten T., Almark U.C.M., Podowski R.M., Naeslund A.K.,  
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;  
RT "The genome sequence of Rickettsia prowazekii and the origin of  
RT mitochondria."  
RL Nature 396:133-140(1998).  
RN [2]  
RP SEQUENCE OF 325-500 FROM N.A.  
RC STRAIN=Madrid E;  
RX MEDLINE=97419517; PubMed=9274032;  
RA Andersson J.O., Andersson S.G.E.;  
RT "Genomic rearrangements during evolution of the obligate  
RT intracellular parasite Rickettsia prowazekii as inferred from an  
RT analysis of 52015 bp nucleotide sequence."  
RL Microbiology 143:2783-2795(1997).  
CC -1- FUNCTION: PROVIDES THE RICKETTSIAL CELL WITH HOST ATP IN  
CC EXCHANGE FOR RICKETTSIAL ADP. THIS IS AN OBLIGATE EXCHANGE SYSTEM.  
CC THIS ENERGY ACQUIRING ACTIVITY IS AN IMPORTANT COMPONENT OF  
CC RICKETTSIAL PARASITISM (BY SIMILARITY)  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- SIMILARITY: BELONGS TO THE ADP/ATP TRANSLOCASE TLC FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AJ235273; CAA15167.1; -  
EMBL; Y11778; CAA72457.1; -  
InterPro; IPR004667; ADP\_ATP\_car.  
Pfam; PF03219; TFC; 1.  
TIGRFAMs; TIGR00769; AAA; 1.  
Transmembrane; Transport; ATP-binding; Multigene family;  
Complete proteome.  
TRANSMEM 26 46 POTENTIAL.  
TRANSMEM 52 82 POTENTIAL.  
TRANSMEM 94 114 POTENTIAL.  
TRANSMEM 149 169 POTENTIAL.  
TRANSMEM 184 204 POTENTIAL.  
TRANSMEM 224 244 POTENTIAL.  
TRANSMEM 287 307 POTENTIAL.  
TRANSMEM 328 348 POTENTIAL.  
TRANSMEM 357 377 POTENTIAL.  
TRANSMEM 381 401 POTENTIAL.  
TRANSMEM 469 489 POTENTIAL.  
SEQUENCE 500 AA; 57073 MW; FE3DB48D08CF5F72 CRC64;

Query Match 55.1%; Score 38; DB 1; Length 500;  
Best Local Similarity 54.5%; Pred. No. 14;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

1 CFQWKRMRKV 11  
||| |:::  
482 CFATYAVRKI 492

SULT 13  
V\_HV2NZ

ENV\_HV2NZ STANDARD; PRT; 856 AA.  
P05883;  
01-NOV-1988 (Rel. 09, Created)  
01-NOV-1988 (Rel. 09, Last sequence update)  
15-JUL-1999 (Rel. 38, Last annotation update)  
Envelope polyprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
ENV.  
Human immunodeficiency virus type 2 (isolate NIH-Z) (HIV-2).  
Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
NCBI\_TaxID=11719;  
[1]  
SEQUENCE FROM N.A.  
MEDLINE=8830359; PubMed=3261862;  
Zagury J.F., Franchini G., Reitz M.S. Jr., Collalti E., Starcich B.R., Hall L., Fargnoli K., Jagodzinski L.L., Guo H.-G., Laure F., Arya S.K., Josephs S.F., Zagury D., Wong-Staal F., Gallo R.C.;  
"Genetic variability between isolates of human immunodeficiency virus (HIV) type 2 is comparable to the variability among HIV type 1.";  
Proc. Natl. Acad. Sci. U.S.A. 85:5941-5945(1988).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; J03654; AAB00761.1; -  
HIV; J03654; ENV2NIH2.  
InterPro; IPR000328; Env GP41.  
InterPro; IPR000777; GP120.  
Pfam; PF00516; GP120; 1.  
Pfam; PF00517; GP41; 1.  
AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;

KW Signal. 20  
FT SIGNAL 21 492 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 493 856 TRANSMEMBRANE GLYCOPROTEIN.  
FT CARBOHYD 37 70 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 70 70 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 79 79 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 112 112 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 116 116 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 128 128 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 133 133 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 182 182 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 183 183 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 196 196 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 228 228 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 231 231 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 268 268 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 279 279 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 290 290 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 355 355 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 440 440 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 457 457 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 602 602 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 627 627 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 856 AA; 99665 MW; A93B0A7E2B881D6 CRC64;

Query Match 55.1%; Score 38; DB 1; Length 856;  
Best Local Similarity 60.0%; Pred. No. 25;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 3 QWKRMRKV 12  
||| |:::  
Db 334 QWKRMRKV 343

RESULT 14

7LES\_DROVI STANDARD; PRT; 2594 AA.  
ID 7LES\_DROVI AC P20806;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Sevenless protein (EC 2.7.1.112).  
GN SEV.  
OS Drosophila virilis (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7244;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90319110; PubMed=2115169;  
RA Michael W.M., Bowtell D.D.L., Rubin G.M.;  
RT "Comparison of the sevenless genes of Drosophila virilis and Drosophila melanogaster.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:5351-5353(1990).  
CC -1- FUNCTION: RECEPTOR FOR AN EXTRACELLULAR SIGNAL REQUIRED TO INSTRUCT A CELL TO DIFFERENTIATE INTO A R7 PHOTORECEPTOR. THE LIGAND FOR SEV IS THE BOSS (BRIDE OF SEVENLESS) PROTEIN.  
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.  
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN RECEPTOR SUBFAMILY.  
CC -1- SIMILARITY: CONTAINS 7 FIBRONECTIN TYPE III-LIKE DOMAINS.  
CC -1- CAUTION: UNCLEAR WHETHER THE POTENTIAL MEMBRANE SPANNING REGION NEAR THE N-TERMINUS IS PRESENT AS A TRANSMEMBRANE DOMAIN IN THE



SEQUENCE 528 AA; 60017 MW; C888DA1391B651EE CRC64;  
Query Match 53.6%; Score 37; DB 1; Length 528;  
Best Local Similarity 50.0%; Pred No 23;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
1 CFQWKRAMRK 10  
|:|:|:|:  
315 CGEWRKPMKR 324

arch completed: February 21, 2003, 07:51:40  
b time : 6.2093 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

protein - protein search, using sw model

on: February 21, 2003, 07:44:43 ; Search time 21.6744 Seconds  
(without alignments)  
114.078 Million cell updates/sec

file: US-09-743-107B-93

fect score: 69

quence: 1 CFQWKRMRKV 12

oring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

arched: 671580 seqs, 206047115 residues

tal number of hits satisfying chosen parameters: 671580

nimum DB seq length: 0

ximum DB seq length: 2000000000

st-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

tabase :

SPREMBL 21:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteriapi:\*

17: sp\_archepa:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

sult No.	Score	Query Match	Length	DB ID	Description
1	54	78.3	711	4	Q8TCD2
2	50	72.5	38	4	Q9UCY5
3	46	66.7	33	6	Q9TR80
4	45	65.2	275	5	Q93780
5	43	62.3	205	16	Q986A0
6	43	62.3	206	16	Q8UHC2
7	43	62.3	206	16	Q92RH8
8	43	62.3	208	16	Q9YFK3
9	42	60.9	1121	10	Q9SW86
10	41	59.4	508	15	Q74118
11	41	59.4	1391	5	O17772
12	40	58.0	105	10	Q9XFD5
13	40	58.0	126	15	Q90827
14	40	58.0	148	10	Q9XHP1
15	40	58.0	274	4	Q96W21
16	40	58.0	830	10	Q945T7

17	40	58.0	1137	10	Q9M7A9
18	40	58.0	1139	10	Q8VWN1
19	39	56.5	57	4	Q96136
20	39	56.5	99	15	P88213
21	39	56.5	154	15	Q9E8H5
22	39	56.5	253	12	Q68541
23	39	56.5	279	16	Q8XSE2
24	39	56.5	397	5	Q9U054
25	39	56.5	453	3	Q04623
26	39	56.5	845	15	Q8UNF5
27	38	55.1	77	15	O91683
28	38	55.1	80	15	O91699
29	38	55.1	91	15	O77855
30	38	55.1	117	15	O73233
31	38	55.1	133	15	O90822
32	38	55.1	207	10	Q9SML1
33	38	55.1	240	10	Q9SML2
34	38	55.1	306	4	Q8TAX2
35	38	55.1	377	12	Q89164
36	38	55.1	377	12	Q85389
37	38	55.1	377	12	Q33122
38	38	55.1	377	12	Q8V2N9
39	38	55.1	377	12	Q8QMT7
40	38	55.1	466	4	Q9NUS2
41	38	55.1	499	10	Q9XFX1
42	38	55.1	659	15	Q9DSL3
43	38	55.1	759	4	Q9BVH6
44	38	55.1	864	5	O62582
45	38	55.1	864	5	Q8SRG3

## ALIGNMENTS

### RESULT 1

Q8TCD2 PRELIMINARY; PRT; 711 AA.  
ID Q8TCD2  
AC Q8TCD2;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Lactotransferin.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PROSTATE;  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BCC22347; AAH22347.1; -  
SQ SEQUENCE 711 AA; 78327 MW; 1B9C7EE097C45FAF CRC64;

Query Match 78.3%; Score 54; DB 4; Length 711;  
Best Local Similarity 81.8%; Pred. No. 0.21;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

### Qy 1 CFQWKRMRKV 11

Db 39 CFQWKRMRKV 49

### RESULT 2

Q9UCY5 PRELIMINARY; PRT; 38 AA.  
ID Q9UCY5  
AC Q9UCY5;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE Lactoferrin homolog (fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

C Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
X NCBI\_TaxID=9606;  
Y [1]  
P SEQUENCE FROM N.A.  
K MEDLINE=96061613; PubMed=8551695;  
S Sato I.;  
I "Characterization of the 84-kDa protein with ABH activity in human  
I seminal plasma.";  
Jpn. J. Legal Med. 49:281-293(1995).  
R HSP; P02788; IBA.  
R InterPro: IPR001156; Transferrin.  
R Pfam: PF00405; transferrin; 1.  
Q SEQUENCE 38 AA; 4459 MW; 040249055BDDDEB CRC64;  
Query Match 72.5%; Score 50; DB 4; Length 38;  
Best Local Similarity 81.8%; Pred. No. 0.058;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Y 2 FQWKRMRKVR 12  
|||:|||||  
C 21 FQWKRMRKVR 31  
|||:|||||  
RESULT 3  
9TR80 PRELIMINARY; PRT; 33 AA.  
D Q9TR80  
I 01-MAY-2000 (TrEMBLrel. 13, Created)  
I 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
I 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
E Lactoferrin (Fragment).  
S Ovis aries (Sheep).  
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
C Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
C Bovidae; Caprinae; Ovis.  
X NCBI\_TaxID=9940;  
Y [1]  
P SEQUENCE.  
K MEDLINE=95127729; PubMed=7827104;  
A Qian Z.Y., Jolles P., Migliore-Samour D., Fiat A.M.;  
L Biochim. Biophys. Acta 1243:25-32(1995).  
R HSP; O77698; ICE2.  
R InterPro: IPR001156; Transferrin.  
R Pfam: PF00405; transferrin; 1.  
Q SEQUENCE 33 AA; 3914 MW; D1904CAB15A73961 CRC64;  
Query Match 66.7%; Score 45; DB 6; Length 33;  
Best Local Similarity 54.5%; Pred. No. 0.26;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
Y 1 CFQWKRMRKVR 11  
|||:|||||  
D 19 CYQWQKMRKL 29  
|||:|||||  
RESULT 4  
93780 PRELIMINARY; PRT; 275 AA.  
D Q93780  
I 01-FEB-1997 (TrEMBLrel. 02, Created)  
I 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
I 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
E F53H4.4 protein.  
N F53H4.4.  
S Caenorhabditis elegans.  
C Eukaryota; Metazoa; Nematozoa; Chromadorea; Rhabditida; Rhabditidae;  
C Rhabditidae; Peloderinae; Caenorhabditis.  
X NCBI\_TaxID=6239;  
Y [1]  
P SEQUENCE FROM N.A.  
A Dobson R.;  
Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
Y [2]

RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C.elegans: A platform for  
RT investigating biology.";  
RL Science 282:2012-2018(1998).  
DR EMBL; Z61089; CAB03137.1; -.  
SQ SEQUENCE 275 AA; 31716 MW; 9D69524FFC704DB2 CRC64;  
Query Match 65.2%; Score 45; DB 5; Length 275;  
Best Local Similarity 72.7%; Pred. No. 3.2;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 2 FQWKRMRKVR 12  
|||:|||||  
DB 262 FQWKRMRKVR 272  
|||:|||||  
RESULT 5  
Q986A0 PRELIMINARY; PRT; 205 AA.  
ID Q986A0  
AC Q986A0  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Pyridoxamine 5'-phosphate oxidase.  
GN MLL7454.  
OS Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Phyllobacteriaceae; Mesorhizobium.  
OX NCBI\_TaxID=381;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MAFF303099;  
RX MEDLINE=21082930; PubMed=11214968;  
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Idekawa K., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,  
RA Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
RT Mesorhizobium loti";  
RL DNA Res. 7:331-338(2000).  
DR EMBL; AF003011; BAB53553.1; -.  
DR InterPro: IPR000659; Pyridox Oxidase.  
DR Pfam: PF01243; Pyridox Oxidase; 1.  
DR ProDom: PD006312; Pyridox Oxidase; 1.  
DR TIGSFams; TIGR00558; pdxH; 1.  
DR PROSITE; PS01064; PYRIDOX\_OXIDASE; 1.  
KW Complete proteome.  
SQ SEQUENCE 205 AA; 23300 MW; 0BADE4CD312327EA CRC64;  
Query Match 62.3%; Score 43; DB 16; Length 205;  
Best Local Similarity 58.3%; Pred. No. 5.5;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 CFQWKRMRKVR 12  
|||:|||||  
DB 88 CFQWKRMRKVR 99  
|||:|||||  
RESULT 6  
Q98UC2 PRELIMINARY; PRT; 206 AA.  
ID Q98UC2  
AC Q98UC2  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Pyridoxamine 5'-phosphate oxidase.  
GN PDXH OR ATU0760 OR AGR\_C1381.  
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Rhizobium.

NCBI\_TaxID=176299;  
[1]  
SEQUENCE FROM N.A.  
MEDLINE=21608550; PubMed=11743193;  
Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,  
Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,  
Chen V., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D.S.,  
Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,  
Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,  
Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,  
Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
Chumley P., Tingey S.V., Tomb J.-P., Gordon M.P., Olson M.V.,  
Nester E.W.;  
"The genome of the natural genetic engineer Agrobacterium tumefaciens  
C58.";  
Science 294:2317-2323 (2001).  
[2]  
SEQUENCE FROM N.A.  
MEDLINE=21608551; PubMed=11743194;  
Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
Quarillo B., Goldman B.S., Cao Y., Askenazi M., Hailing C., Mullin L.,  
Houmel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,  
William C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,  
Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,  
Cielo C., Slater S.;  
"Genome sequence of the plant pathogen and biotechnology agent  
Agrobacterium tumefaciens C58.";  
Science 294:2323-2328 (2001).  
ENBL; AG009043; AAL41776.1; -;  
ENBL; AG008009; AAK8569.1; -;  
Complete proteome.  
SEQUENCE 206 AA; 23720 MW; 3BE488AE5307C0C1 CRC64;  
Query Match 62.3%; Score 43; DB 16; Length 206;  
Best Local Similarity 58.3%; Pred. No. 5.5; Mismatches 4; Indels 0; Gaps 0;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
1 CFQWKEAMRKVR 12  
|||  
88 CFHWKSLRRQVR 99  
[1]  
[1]  
ULT 7  
RH8  
Q92RH8 PRELIMINARY; PRT; 206 AA.  
01-DEC-2001 (TrEMBLrel. 19, Created)  
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
Probable pyridoxamine 5'-phosphate oxidase (PNP/PPM oxidase) protein  
(EC 1.4.3.5).  
PDXH OR R00895 OR SMC00069.  
Rhizobium meliloti (Sinorhizobium meliloti).  
Bacteria; Proteobacteria; alpha subphylum; Rhizobiaceae group;  
Rhizobiaceae; Sinorhizobium.  
NCBI\_TaxID=382;  
[1]  
SEQUENCE FROM N.A.  
STRAIN=1021;  
MEDLINE=21396507; PubMed=11481430;  
Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
Boisard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,  
Pohl T., Portetelle D., Puehler A., Purnelle B., Rampsberger U.,  
Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;  
"Analysis of the chromosome sequence of the legume symbiont  
Sinorhizobium meliloti strain 1021.";  
Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882 (2001).  
ENBL; AL591785; CAC45467.1; -;  
InterPro; IPR000659; Pyridox\_oxidase.  
Pfam; PF01243; Pyridox\_oxidase; 1.  
ProDom; PD006312; Pyridox\_oxidase; 1.

DR TIGRFAMs; TIGR00558; pdxH; 1.  
DR PROSITE; PS01064; PYRIDOX\_OXIDASE; UNKNOWN\_1.  
KW Oxidoreductase; Complete proteome.  
SQ SEQUENCE 206 AA; 23900 MW; A2DB74229DAC97A CRC64;  
Query Match 62.3%; Score 43; DB 16; Length 206;  
Best Local Similarity 58.3%; Pred. No. 5.5; Mismatches 4; Indels 0; Gaps 0;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
1 CFQWKEAMRKVR 12  
|||  
88 CFHWKSLRRQVR 99  
[1]  
[1]  
ULT 8  
Q8YFK3 PRELIMINARY; PRT; 208 AA.  
AC Q8YFK3;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Probable pyridoxamine 5'-phosphate oxidase (EC 1.4.3.5).  
ENBL; I11517.  
GN Brucella melitensis.  
OC Bacteria; Proteobacteria; alpha subphylum; Rhizobiaceae group;  
OC Brucellaceae; Brucella.  
NCBI\_TaxID=29459;  
RN [1]  
RP SEQUENCE FROM N.A.  
STRAIN=16M / ATCC 23456 / BIOTYPE 1;  
RC MEDLINE=20020109; PubMed=11756688;  
RA DelVecchio V.G., Kapral V., Redkar R.J., Patra G., Mujar C., Los T.,  
Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,  
Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,  
Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,  
Haselkorn R., Kyrides N., Overbeek R.;  
"The genome sequence of the facultative intracellular pathogen  
Brucella melitensis.";  
Proc. Natl. Acad. Sci. U.S.A. 99:443-448 (2002).  
RL ENBL; AG009587; AAL52698.1; -;  
DR InterPro; IPR000659; Pyridox\_oxidase.  
DR Pfam; PF01243; Pyridox\_oxidase; 1.  
DR ProDom; PD006312; Pyridox\_oxidase; 1.  
DR TIGRFAMs; TIGR00558; pdxH; 1.  
DR PROSITE; PS01064; PYRIDOX\_OXIDASE; 1.  
KW Oxidoreductase; Complete proteome.  
SQ SEQUENCE 208 AA; 23866 MW; CB1F50BC9612DE28 CRC64;  
Query Match 62.3%; Score 43; DB 16; Length 208;  
Best Local Similarity 58.3%; Pred. No. 5.5; Mismatches 4; Indels 0; Gaps 0;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
1 CFQWKEAMRKVR 12  
|||  
90 CFHWKSLRRQVR 101  
[1]  
[1]  
ULT 9  
Q9SWS6 PRELIMINARY; PRT; 1121 AA.  
AC Q9SWS6;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Phycochrome B2.  
GN PHYB2.  
OS Lycopersicon esculentum (Tomato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
NCBI\_TaxID=4081;  
RN [1]  
RP SEQUENCE FROM N.A.

```

MEDLINE=99413290; PubMed=10485280;
Kerckhoffs L.H., Kelmenson P.M., Schreuder M.E., Kendrick C.I.,
Kendrick R.E., Hanhart C.J., Koorneef M., Pratt L.H.,
Cordonnier-Pratt M.M.;
"Characterization of the gene encoding the apoprotein of phytochrome
B2 in tomato, and identification of molecular lesions in two mutant
alleles.";
Mol. Gen. Genet. 261:901-907(1999).
EMBL; AF122901; AAD50631.1; -.
InterPro; IPR001298; ADH short.
InterPro; IPR003594; ATPBind_ATPase.
InterPro; IPR003018; GAF.
InterPro; IPR003661; His_kinA.
InterPro; IPR004359; His_KIN_sig.
InterPro; IPR001610; PAC.
InterPro; IPR000014; PAS domain.
InterPro; IPR001294; Phytochrome.
Pfam; PF01590; GAF; 1.
Pfam; PF02518; HATPase_c; 1.
Pfam; PF00389; PAS; 2.
Pfam; PF00360; phytochrome; 1.
Pfam; PF00512; signal; 1.
PRINTS; PR01033; PHYTOCHROME.
SMART; SM00065; GAF; 1.
SMART; SM00387; HATPase_c; 1.
SMART; SM00388; HisK_A; 1.
SMART; SM00086; PAC; 1.
SMART; SM00091; PAS; 2.
TIGRFAMs; TIGR00229; sensory_box; 2.
PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
PROSITE; PS00245; PHYTOCHROME_1; 1.
PROSITE; PS00046; PHYTOCHROME_2; 1.
SEQUENCE 1121 AA; 125308 MW; ED9EDA704BB37F27 CRC64;

Query Match 60.9%; Score 42; DB 10; Length 1121;
Best Local Similarity 54.5%; Pred. No. 44;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

1 CFQWKRAVRKVR 11
|||:|:|:|
777 CFENWTAMEKL 787

RESULT 10
74118 PRELIMINARY; PRT; 508 AA.
Q74118;
01-NOV-1996 (TREMBlrel. 01, Created)
01-NOV-1996 (TREMBlrel. 01, Last sequence update)
01-DEC-2001 (TREMBlrel. 19, Last annotation update)
Gp105 (Fragment).
ENV.
Human immunodeficiency virus type 2.
Viruses; Retroid viruses; Retroviridae; Lentivirus.
NCBI_TaxID=11709;
[1] _
SEQUENCE FROM N.A.
STRAIN=HIV2D868;
MEDLINE=94187055; PubMed=8139000;
Grez M., Dietrich U., Balfe P., Von Briesen H., Maniar J.K.,
Mahabre G., Delwart E.L., Mullins J.I., Ruebamen-Waigmann H.;
"Genetic analysis of Human immunodeficiency virus type 1 and 2 (HIV-1
and HIV-2) mixed infections in India reveals a recent spread of HIV-1
and HIV-2 from a single ancestor for each of these viruses.";
J. Virol. 68:2161-2168(1994).
EMBL; U07108; AAA17669.1; -.
InterPro; IPR000777; GP120.
Pfam; PF00516; GP120; 1.
AIDS; Coat protein; Glycoprotein.
NCBI_TaxID=508;
NON_TER 508 AA; 58165 MW; 9B103784A08DCB9C CRC64;

Query Match 59.4%; Score 41; DB 15; Length 508;

```

```

Best Local Similarity 70.0%; Pred. No. 30;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 QWKRAVRKVR 12
:|||||:|
DB 340 KWKEARREVR 349

RESULT 11
017772 PRELIMINARY; PRT; 1391 AA.
ID 017772;
AC 017772;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE F09C3.1 protein.
GN F09C3.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peleoderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Lennard N.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z92781; CAB07179.2; -.
DR InterPro; IPR001715; Galponin-like.
DR InterPro; IPR000048; IQ region.
DR InterPro; IPR001936; RasGAP.
DR InterPro; IPR000593; RasGAP_C.
DR Pfam; PF00612; IQ; 1.
DR ProDom; PD008735; RasGAP_C; 1.
DR SMART; SM00033; CH; 1.
DR SMART; SM00015; IQ; 1.
DR PROSITE; PS50021; CH; 1.
DR PROSITE; PS50018; RAS_GTPASE_ACTIV_2; 1.
SQ SEQUENCE 1391 AA; 159210 MW; 794DF9EBFA2B65AC CRC64;

Query Match 59.4%; Score 41; DB 5; Length 1391;
Best Local Similarity 70.0%; Pred. No. 83;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 QWKRAVRKVR 12
:|||||:|
DB 122 QWRRAVESVR 131

RESULT 12
Q9XFD5 PRELIMINARY; PRT; 105 AA.
ID Q9XFD5;
AC Q9XFD5;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Cytochrome P450 (Fragment).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartioideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1] _
RP SEQUENCE FROM N.A.
RA TISSUE=PANICLE;
RA Liu J., Yang J.;
RT "Suppression subtractive hybridization (SSH) identified candidate
genes that are differentially expressed at rice young panicle.";

```



Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.

EMBL; AF140486; AAD29699.1; -  
 InterPro; IPR001128; Cytochrome\_P450.

Pfam; PF00067; P450; 1.  
 PRINTS; PR00385; P450.

PROSITE; PS00086; CYTOCHROME\_P450; UNKNOWN\_1.  
 Heme; Monooxygenase; Oxidoreductase.

NON\_TER 1  
 SEQUENCE 105 AA; 11912 MW; B0EFCDD487E19F9 CRC64;

Query Match 58.0%; Score 40; DB 10; Length 105;  
 Best Local Similarity 60.0%; Pred. No. 9.6;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

1 CFQWKRAMRK 10  
 |||||:|

61 CFQWERLGRK 70  
 |||||:|

RESULT 13

090827 PRELIMINARY; PRT; 126 AA.

090827; (TrEMBLrel. 08, Created)

01-NOV-1998 (TrEMBLrel. 08, Last sequence update)

01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

Envelope protein (Fragment).

ENV.

Human immunodeficiency virus type 2.

Viruses; Retroviral viruses; Retroviridae; Lentivirus.

NCBI\_TaxID=11709;

SEQUENCE FROM N.A.

STRAIN-A; MEDLINE=97255645; PubMed=9100992;

Xiang Z., Ariyoshi K., Wilkins A., Dias F., Whittle H., Breuer J.;  
 "HIV type 2 pathogenicity is not related to subtype in rural Guinea  
 bissau."

AIDS Res. Hum. Retroviruses 13:501-505(1997).

SEQUENCE FROM N.A.

STRAIN-A; MEDLINE=98406190; PubMed=9733826;

Grassly N., Xiang Z., Ariyoshi K., Aaby P., Jensen H., Dias F.,  
 Van der Loeff, Whittle H., Breuer J.;

"Mortality among human immunodeficiency virus type 2-positive  
 villagers in rural Guinea-bissau is correlated with viral genotype."

J. Virol. 72:7893-7899(1998).

EMBL; AJ011269; CAA09580.1; -

InterPro; IPR000777; GPI20.

Pfam; PF00516; GPI20; 1.

AIDS; Coat protein; Glycoprotein.

NON\_TER 1  
 NON\_TER 126 126

SEQUENCE 126 AA; 14891 MW; 248828768F66F888 CRC64;

Query Match 58.0%; Score 40; DB 15; Length 126;  
 Best Local Similarity 60.0%; Pred. No. 12;

Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

3 QWKRAMRKVR 12  
 |||||:|

72 EWRKAMQEVK 81  
 |||||:|

RESULT 14

IXHP1

09XHP1 PRELIMINARY; PRT; 148 AA.

09XHP1; (TrEMBLrel. 12, Created)

01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

01-NAR-2002 (TrEMBLrel. 20, Last annotation update)

DE 2S albumin.

OS Sesamum indicum (Oriental sesame) (Gingelly).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; euasterids I; Lamiales; Pedaliaceae; Sesamum.

OX NCBI\_TaxID=4182;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-TAINAN 1;

RX MEDLINE=20074970; PubMed=10606554;

RA Tai S.S.K., Wu L.S.H., Chen E.C.F., Tzen J.T.C.;

"Molecular cloning of 11S globulin and 2S albumin, the two major seed  
 storage proteins in sesame."

J. Agric. Food Chem. 47:4932-4938(1999).

DR EMBL; AF091841; AAD42943.1; -

DR InterPro; IPR003612; AAI.

DR InterPro; IPR00617; Napin.

DR Pfam; PF00234; tryp\_alpha\_aml; 1.

DR PRINTS; PR00496; NAFIN.

DR SMART; SM00499; AAI; 1.

SQ SEQUENCE 148 AA; 17524 MW; BA46B033BA13E3DE CRC64;

Query Match 58.0%; Score 40; DB 10; Length 148;  
 Best Local Similarity 66.7%; Pred. No. 14;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWKRAMR 9  
 |||||:|

Db 54 CQWVRSMR 62  
 |||||:|

RESULT 15

Q96M21

ID Q96M21 PRELIMINARY; PRT; 274 AA.

AC Q96M21;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE CDNA FLJ32891 fis, clone TESTI2004929.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,

RA Arita M., Musashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,

RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,

RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,

RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,

RA Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,

RA Sugano S., Nagahori K., Masuho Y., Nagai K., Isogai T.;

"NDO human cDNA sequencing project."

RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

RL EMBL; AK057453; BAB71493.1; -

SQ SEQUENCE 274 AA; 30083 MW; 1DD43654D4135B2F CRC64;

Query Match 58.0%; Score 40; DB 4; Length 274;  
 Best Local Similarity 50.0%; Pred. No. 25;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWKRAMRKVR 12  
 |||||:|

Db 66 CFQWVGVRVYLR 77  
 |||||:|

Search completed: February 21, 2003, 08:00:46  
 Job time : 22.6744 secs

GenCore version 5.1.3  
Copyright (C) 1993 - 2003 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: February 21, 2003, 07:37:21 ; Search time 28.093 Seconds  
(without alignments)  
56.918 Million cell updates/sec

Title: 'us-09-743-107b-94  
Perfect score: 70  
Sequence: 1 CFAWRNMRKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq 101002:\*  
1: /SID32/gcgdata/geneseq/geneseq-emb1/AA1980.DAT\*  
2: /SID32/gcgdata/geneseq/geneseq-emb1/AA1981.DAT\*  
3: /SID32/gcgdata/geneseq/geneseq-emb1/AA1982.DAT\*  
4: /SID32/gcgdata/geneseq/geneseq-emb1/AA1983.DAT\*  
5: /SID32/gcgdata/geneseq/geneseq-emb1/AA1984.DAT\*  
6: /SID32/gcgdata/geneseq/geneseq-emb1/AA1985.DAT\*  
7: /SID32/gcgdata/geneseq/geneseq-emb1/AA1986.DAT\*  
8: /SID32/gcgdata/geneseq/geneseq-emb1/AA1987.DAT\*  
9: /SID32/gcgdata/geneseq/geneseq-emb1/AA1988.DAT\*  
10: /SID32/gcgdata/geneseq/geneseq-emb1/AA1989.DAT\*  
11: /SID32/gcgdata/geneseq/geneseq-emb1/AA1990.DAT\*  
12: /SID32/gcgdata/geneseq/geneseq-emb1/AA1991.DAT\*  
13: /SID32/gcgdata/geneseq/geneseq-emb1/AA1992.DAT\*  
14: /SID32/gcgdata/geneseq/geneseq-emb1/AA1993.DAT\*  
15: /SID32/gcgdata/geneseq/geneseq-emb1/AA1994.DAT\*  
16: /SID32/gcgdata/geneseq/geneseq-emb1/AA1995.DAT\*  
17: /SID32/gcgdata/geneseq/geneseq-emb1/AA1996.DAT\*  
18: /SID32/gcgdata/geneseq/geneseq-emb1/AA1997.DAT\*  
19: /SID32/gcgdata/geneseq/geneseq-emb1/AA1998.DAT\*  
20: /SID32/gcgdata/geneseq/geneseq-emb1/AA1999.DAT\*  
21: /SID32/gcgdata/geneseq/geneseq-emb1/AA2000.DAT\*  
22: /SID32/gcgdata/geneseq/geneseq-emb1/AA2001.DAT\*  
23: /SID32/gcgdata/geneseq/geneseq-emb1/AA2002.DAT\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	70	100.0	12	21	AAV78094 Human lactoferrin
2	66	94.3	12	21	AAV78070 Human lactoferrin
3	65	92.9	12	21	AAV78091 Human lactoferrin
4	65	92.9	12	21	AAV78092 Human lactoferrin
5	61	87.1	12	21	AAV78038 Human lactoferrin
6	61	87.1	12	21	AAV78046 Human lactoferrin
7	61	87.1	12	21	AAV78047 Human lactoferrin
8	61	87.1	12	21	AAV78084 Human lactoferrin
9	61	87.1	12	21	AAV78037 Human lactoferrin
10	61	87.1	13	21	AAV78048 Human lactoferrin

11	61	87.1	13	21	AAV78049 Human lactoferrin
12	61	87.1	14	21	AAV78036 Human lactoferrin
13	61	87.1	14	21	AAV78050 Human lactoferrin
14	61	87.1	14	21	AAV78051 Human lactoferrin
15	61	87.1	15	17	AAV78054 Peptide for anti-u
16	61	87.1	15	21	AAV78035 Human lactoferrin
17	61	87.1	15	21	AAV78062 Human lactoferrin
18	61	87.1	15	21	AAV78063 Human lactoferrin
19	61	87.1	16	21	AAV78031 Human lactoferrin
20	61	87.1	16	21	AAV78064 Human lactoferrin
21	61	87.1	16	21	AAV78065 Human lactoferrin
22	61	87.1	17	21	AAV78034 Human lactoferrin
23	61	87.1	17	21	AAV78066 Human lactoferrin
24	61	87.1	17	21	AAV78067 Human lactoferrin
25	61	87.1	18	15	AAV69352 Human lactoferrin
26	61	87.1	18	17	AAV13397 Advanced glycosyla
27	61	87.1	18	21	AAV78033 Human lactoferrin
28	61	87.1	19	21	AAV68867 Amino acid sequenc
29	61	87.1	19	21	AAV78032 Human lactoferrin
30	61	87.1	20	13	AAV21810 Anti microbial pep
31	61	87.1	20	14	AAV44841 Lactoferrin-relate
32	61	87.1	20	15	AAV48530 Lactoferrin derive
33	61	87.1	20	15	AAV48531 Lactoferrin derive
34	61	87.1	20	15	AAV57461 Lactoferrin derive
35	61	87.1	20	15	AAV57462 Lactoferrin derive
36	61	87.1	20	16	AAV84698 Bovine lactoferrin
37	61	87.1	20	16	AAV84699 Bovine lactoferrin
38	61	87.1	20	16	AAV80263 Anti-parasitic lac
39	61	87.1	20	16	AAV80264 Anti-parasitic lac
40	61	87.1	20	17	AAV98553 Peptide for anti-u
41	61	87.1	20	17	AAV98552 Lactoferrin-derive
42	61	87.1	20	17	AAV03045 Lactoferrin-derive
43	61	87.1	20	17	AAV90607 Lactoferrin-derive
44	61	87.1	20	17	AAV87621 Lactoferrin-derive
45	61	87.1	20	17	AAV87622 Lactoferrin-derive

## ALIGNMENTS

RESULT 1  
AAV78094  
ID AAV78094 standard; Peptide; 12 AA.  
XX AAV78094;  
XX 25-APR-2000 (first entry)  
XX Human lactoferrin derived peptide SEQ ID NO:94.  
XX Human; lactoferrin; modification; infection; inflammation; tumour;  
XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
XX urinary tract infection; colitis; Candida infection; fungicidal;  
XX bactericidal; preservative.  
XX Homo sapiens.  
XX OS Synthetic.  
XX WO200001730-A1.  
XX 13-JAN-2000.  
XX 06-JUL-1999; 99WO-SE01230.  
XX 06-JUL-1998; 98SE-0002441.  
XX 17-JUL-1998; 98SE-0002562.  
XX 29-DEC-1998; 98SE-0004614.  
XX (ASCI-) A+ SCI INVEST AB.  
XX Hanson LA, Matteby-Baltzer I, Baltzer L, Dolphin GT;  
XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
XX  
PS Claim 22; Page 38; 102pp; English.  
XX  
CC AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.  
XX  
XX Sequence 12 AA;  
SQ Query Match 100.0%; Score 70; DB 21; Length 12;  
Best Local Similarity 100.0%; Pred. No. 3e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 CFAWKRNMRKVR 12  
DB 1 CFAWKRNMRKVR 12  
|||:|||||  
RESULT 2  
ID AAY78070 standard; Peptide; 12 AA.  
XX AAY78070;  
AC AAY78070;  
XX 25-APR-2000 (first entry)  
XX Human lactoferrin derived peptide SEQ ID NO:70.  
DE Human; lactoferrin; modification; infection; inflammation; tumour;  
XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
XX bactericidal; preservative.  
XX Homo sapiens.  
OS Synthetic.  
XX WO200001730-A1.  
PN 13-JAN-2000.  
XX 06-JUL-1999; 99WO-SE01230.  
XX 06-JUL-1998; 98SE-0002441.  
XX 17-JUL-1998; 98SE-0002562.  
XX 29-DEC-1998; 98SE-0004614.  
XX (ASCI-) A+ SCI INVEST AB.  
XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX WPI; 2000-147388/13.  
XX New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
XX  
XX Claim 22; Page 35; 102pp; English.  
XX  
XX AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.  
XX  
XX Sequence 12 AA;  
SQ Query Match 94.3%; Score 66; DB 21; Length 12;  
Best Local Similarity 91.7%; Pred. No. 0.00014;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
OY 1 CFAWKRNMRKVR 12  
DB 1 CFAWKRNMRKVR 12  
|||:|||||  
RESULT 3  
ID AAY78091 standard; Peptide; 12 AA.  
XX AAY78091;  
AC AAY78091;  
XX 25-APR-2000 (first entry)  
XX Human lactoferrin derived peptide SEQ ID NO:91.  
DE Human; lactoferrin; modification; infection; inflammation; tumour;  
XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
XX bactericidal; preservative.  
XX Homo sapiens.  
OS Synthetic.  
XX WO200001730-A1.  
PN 13-JAN-2000.  
XX 06-JUL-1999; 99WO-SE01230.  
XX 06-JUL-1998; 98SE-0002441.  
XX 17-JUL-1998; 98SE-0002562.  
XX 29-DEC-1998; 98SE-0004614.  
XX (ASCI-) A+ SCI INVEST AB.  
XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX WPI; 2000-147388/13.  
XX New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
XX  
XX Claim 22; Page 38; 102pp; English.  
XX  
XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also

AA78001 to AA78100 represent peptides having sequences based on human lactoferrin. The peptides are taken up in the intestine through binding to specific lactoferrin receptors and are then transported through the circulation. A medicinal product of the peptide or fragment can be used for treating and/or prevention of infections (such as urinary tract infections, colitis, and Candida infection on a mucosal membrane), inflammations and/or tumours. The peptides can also be used in food stuffs such as infant formula food. The peptides are also fungicidal and bactericidal and may also be used as preservatives. Even though native human lactoferrin have been shown to have desired anti-inflammatory anti-infectious and anti-tumoural properties they cannot be used clinically on a broad basis because of high production costs. Therefore, provision of peptides based on lactoferrin would enable them to be used for the same purposes as lactoferrin at lower cost.

QY 1 CFAWKNNRKVR 12  
DB 1 CFQWQNNRKVR 12

## RESULT 6

AAV78046  
ID AAY78046 standard; Peptide; 12 AA.  
XX AAY78046;  
AC AAY78046;  
XX 25-APR-2000 (first entry)  
DT Human lactoferrin derived peptide SEQ ID NO:46.  
DE Human; lactoferrin; modification; infection; inflammation; tumour;  
XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX Homo sapiens.  
OS Synthetic.  
XX WO200001730-A1.  
PN 13-JAN-2000.  
PD 06-JUL-1999; 99WO-SE01230.  
XX 06-JUL-1998; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
XX (ASCII-) A+ SCI INVEST AB.  
PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX WPI; 2000-147388/13.  
DR New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
PS Claim 15; Page 35; 102pp; English.  
XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumors. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.

XX Sequence 12 AA;  
SQ

Query Match 87.1%; Score 61; DB 21; Length 12;  
Best Local Similarity 83.3%; Pred. NO. 0.00095;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFAWKNNRKVR 12  
DB 1 CFQWQNNRKVR 12

## RESULT 7

AAV78047

ID AAY78047 standard; Peptide; 12 AA.  
XX AAY78047;  
AC AAY78047;  
XX 25-APR-2000 (first entry)  
DT Human lactoferrin derived peptide SEQ ID NO:47.  
DE Human; lactoferrin; modification; infection; inflammation; tumour;  
XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX Homo sapiens.  
OS Synthetic.  
XX WO200001730-A1.  
PN 13-JAN-2000.  
PD 06-JUL-1999; 99WO-SE01230.  
XX 06-JUL-1998; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
XX (ASCII-) A+ SCI INVEST AB.  
PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX WPI; 2000-147388/13.  
DR New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
PS Claim 18; Page 73; 102pp; English.  
XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumors. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.

XX Sequence 12 AA;  
SQ

Query Match 87.1%; Score 61; DB 21; Length 12;  
Best Local Similarity 83.3%; Pred. NO. 0.00095;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFAWKNNRKVR 12  
DB 1 CFQWQNNRKVR 12

## RESULT 8

AAV78084

ID AAY78084 standard; Peptide; 12 AA.  
XX AAY78084;  
AC AAY78084;  
XX 25-APR-2000 (first entry)  
DT Human lactoferrin derived peptide SEQ ID NO:84.  
DE

```

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;
XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;
XX KW urinary tract infection; colitis; Candida infection; fungicidal;
XX KW bactericidal; preservative.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN WO200001730-A1.
XX PD 13-JAN-2000.
XX PF 06-JUL-1999; 99WO-SE01230.
XX PR 06-JUL-1998; 98SE-0002441.
XX PR 17-JUL-1998; 98SE-0002562.
XX PR 29-DEC-1998; 98SE-0004614.
XX PA (ASCI-) A+ SCI INVEST AB.
XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;
XX PF WIPI; 2000-147388/13.
XX PT New peptides used for treatment and prevention of infections,
XX PT inflammations and tumors and for use in infant formula food -
XX PS Claim 22; Page 36; 102pp; English.
XX CC AAY78001 to AAY78100 represent peptides having sequences based on human
XX CC lactoferrin. The peptides are taken up in the intestine through
XX CC binding to specific lactoferrin receptors and are then transported
XX CC through the circulation. A medicinal product of the peptide or fragment
XX CC can be used for treating and/or prevention of infections (such as
XX CC urinary tract infections, colitis, and Candida infection on a mucosal
XX CC membrane), inflammations and/or tumours. The peptides can also be used
XX CC in food stuffs such as infant formula food. The peptides are also
XX CC fungicidal and bactericidal and may also be used as preservatives.
XX CC Even though native human lactoferrin have been shown to have desired
XX CC anti-inflammatory anti-infectious and anti-tumoural properties they
XX CC cannot be used clinically on a broad basis because of high production
XX CC costs. Therefore, provision of peptides based on lactoferrin would
XX CC enable them to be used for the same purposes as lactoferrin at lower
XX CC cost.
XX SQ Sequence 12 AA;
Query Match 87.1%; Score 61; DB 21; Length 12;
Best Local Similarity 83.3%; Pred. No. 0.00095;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 CFAWKNNMKVR 12
DB 1 CFQWQNNMKVR 12
RESULT 9
AAY78037
ID AAY78037 standard; Peptide; 13 AA.
AC AAY78037;
XX 25-APR-2000 (first entry)
XX Human lactoferrin derived peptide SEQ ID NO:37.
XX Human; lactoferrin; modification; infection; inflammation; tumour;
XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;
XX KW urinary tract infection; colitis; Candida infection; fungicidal;
XX KW bactericidal; preservative.
XX OS Homo sapiens.

```

```

OS Synthetic.
XX WO200001730-A1.
XX PD 13-JAN-2000.
XX PF 06-JUL-1999; 99WO-SE01230.
XX PR 06-JUL-1998; 98SE-0002441.
XX PR 17-JUL-1998; 98SE-0002562.
XX PR 29-DEC-1998; 98SE-0004614.
XX PA (ASCI-) A+ SCI INVEST AB.
XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;
XX PF WIPI; 2000-147388/13.
XX PT New peptides used for treatment and prevention of infections,
XX PT inflammations and tumors and for use in infant formula food -
XX PS Claim 12; Page 70; 102pp; English.
XX CC AAY78001 to AAY78100 represent peptides having sequences based on human
XX CC lactoferrin. The peptides are taken up in the intestine through
XX CC binding to specific lactoferrin receptors and are then transported
XX CC through the circulation. A medicinal product of the peptide or fragment
XX CC can be used for treating and/or prevention of infections (such as
XX CC urinary tract infections, colitis, and Candida infection on a mucosal
XX CC membrane), inflammations and/or tumours. The peptides can also be used
XX CC in food stuffs such as infant formula food. The peptides are also
XX CC fungicidal and bactericidal and may also be used as preservatives.
XX CC Even though native human lactoferrin have been shown to have desired
XX CC anti-inflammatory anti-infectious and anti-tumoural properties they
XX CC cannot be used clinically on a broad basis because of high production
XX CC costs. Therefore, provision of peptides based on lactoferrin would
XX CC enable them to be used for the same purposes as lactoferrin at lower
XX CC cost.
XX SQ Sequence 13 AA;
Query Match 87.1%; Score 61; DB 21; Length 13;
Best Local Similarity 83.3%; Pred. No. 0.001;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 CFAWKNNMKVR 12
DB 2 CFQWQNNMKVR 13
RESULT 10
AAY78048
ID AAY78048 standard; Peptide; 13 AA.
AC AAY78048;
XX 25-APR-2000 (first entry)
XX Human lactoferrin derived peptide SEQ ID NO:48.
XX Human; lactoferrin; modification; infection; inflammation; tumour;
XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;
XX KW urinary tract infection; colitis; Candida infection; fungicidal;
XX KW bactericidal; preservative.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN WO200001730-A1.
XX PD 13-JAN-2000.
XX PF 06-JUL-1999; 99WO-SE01230.

```

XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 PA (ASCI-) A+ SCI INVEST AB.

PI Hanson LA, Mattesby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 15; Page 74; 102pp; English.

CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 13 AA;

Query Match 87.1%; Score 61; DB 21; Length 13;  
 Best Local Similarity 83.3%; Pred. No. 0.001;  
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFAWKRNMRKVR 12  
 DB 2 CFQWRNMRKVR 13

RESULT 11  
 AAY78049  
 ID AAY78049 standard; Peptide; 13 AA.

XX AAY78049;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:49.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.  
 OS Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

PI Hanson LA, Mattesby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 18; Page 74; 102pp; English.

CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 13 AA;

Query Match 87.1%; Score 61; DB 21; Length 13;  
 Best Local Similarity 83.3%; Pred. No. 0.001;  
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFAWKRNMRKVR 12  
 DB 2 CFQWRNMRKVR 13

RESULT 12  
 AAY78036  
 ID AAY78036 standard; Peptide; 14 AA.

XX AAY78036;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:36.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.  
 OS Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

PI Hanson LA, Mattesby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

PS Claim 12; Page 69; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through

CC binding to specific lactoferrin receptors and are then transported

CC through the circulation. A medicinal product of the peptide or fragment

CC can be used for treating and/or prevention of infections (such as

CC urinary tract infections, colitis, and Candida infection on a mucosal

CC membrane), inflammations and/or tumours. The peptides can also be used

CC in food stuffs such as infant formula food. The peptides are also

CC even though native human lactoferrin have been shown to have desired

CC anti-inflammatory anti-infectious and anti-tumoural properties they

CC cannot be used clinically on a broad basis because of high production

CC costs. Therefore, provision of peptides based on lactoferrin would

CC enable them to be used for the same purposes as lactoferrin at lower

CC cost.

XX Sequence 14 AA;

SQ Query Match 87.1%; Score 61; DB 21; Length 14;

Best Local Similarity 83.3%; Pred. No. 0.0011;

Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQNNMKVR 12

DB 3 CFQWQNNMKVR 14

RESULT 13

AAY78050

ID AAY78050 standard; Peptide; 14 AA.

XX AC AAY78050;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:50.

XX Human; lactoferrin; modification; infection; inflammation; tumour;

KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;

KW urinary tract infection; colitis; Candida infection; fungicidal;

KW bactericidal; preservative.

XX OS Homo sapiens.

OS Synthetic.

PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX DR New peptides used for treatment and prevention of infections,

PT inflammations and tumors and for use in infant formula food -

XX Claim 15; Page 75; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through

CC binding to specific lactoferrin receptors and are then transported

CC through the circulation. A medicinal product of the peptide or fragment

CC can be used for treating and/or prevention of infections (such as

CC urinary tract infections, colitis, and Candida infection on a mucosal

CC membrane), inflammations and/or tumours. The peptides can also be used

CC in food stuffs such as infant formula food. The peptides are also

CC even though native human lactoferrin have been shown to have desired

CC anti-inflammatory anti-infectious and anti-tumoural properties they

CC cannot be used clinically on a broad basis because of high production

CC costs. Therefore, provision of peptides based on lactoferrin would

CC enable them to be used for the same purposes as lactoferrin at lower

CC cost.

CC urinary tract infections, colitis, and Candida infection on a mucosal

CC membrane), inflammations and/or tumours. The peptides can also be used

CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.

CC Even though native human lactoferrin have been shown to have desired

CC anti-inflammatory anti-infectious and anti-tumoural properties they

CC cannot be used clinically on a broad basis because of high production

CC costs. Therefore, provision of peptides based on lactoferrin would

CC enable them to be used for the same purposes as lactoferrin at lower

CC cost.

XX Sequence 14 AA;

SQ Query Match 87.1%; Score 61; DB 21; Length 14;

Best Local Similarity 83.3%; Pred. No. 0.0011;

Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQNNMKVR 12

DB 3 CFQWQNNMKVR 14

RESULT 14

AAY78051

ID AAY78051 standard; Peptide; 14 AA.

XX AC AAY78051;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:51.

XX Human; lactoferrin; modification; infection; inflammation; tumour;

KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;

KW urinary tract infection; colitis; Candida infection; fungicidal;

KW bactericidal; preservative.

XX OS Homo sapiens.

OS Synthetic.

PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX DR New peptides used for treatment and prevention of infections,

PT inflammations and tumors and for use in infant formula food -

XX Claim 18; Page 75; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through

CC binding to specific lactoferrin receptors and are then transported

CC through the circulation. A medicinal product of the peptide or fragment

CC can be used for treating and/or prevention of infections (such as

CC urinary tract infections, colitis, and Candida infection on a mucosal

CC membrane), inflammations and/or tumours. The peptides can also be used

CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.

CC Even though native human lactoferrin have been shown to have desired

CC anti-inflammatory anti-infectious and anti-tumoural properties they

CC cannot be used clinically on a broad basis because of high production

CC cost.



CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX  
 XX  
 SQ Sequence 14 AA;  
 Query Match 87.1%; Score 61; DB 21; Length 14;  
 Best Local Similarity 83.3%; Pred. No. 0.0011;  
 Matches 10; Conservative 1; Mismatches 0; Gaps 0;

QY 1 CFAWKRNMRKVR 12  
 |||:|||||  
 Db 3 CFQWQRNMRKVR 14

## RESULT 15

AAR98554  
 ID AAR98554 standard; Peptide; 15 AA.

AC AAR98554;

DT 12-NOV-1996 (first entry)

DE Peptide for anti-ulcer agent.

KW anti-ulcer agent; low toxicity; stable; heat-resistant.

OS Synthetic.

FN JP08143468-A.

PD 04-JUN-1996.

PF 17-NOV-1994; 94JP-0283869.

PR 17-NOV-1994; 94JP-0283869.

PA (MORG) MORINAGA MILK IND CO LTD.

DR WPI; 1996-318857/32.

PT Anti-ulcer agent contg. peptide - has low toxicity, is  
 PT heat-resistant and water-soluble

PS Claim 1; Page 11; 11pp; Japanese.

CC AAR98531-54 are peptides used in an anti-ulcer agent. The agent is low  
 CC in toxicity, is heat-resistant and stable in aqueous soln.. It can be  
 CC administered orally and be produced in large amounts.

SQ Sequence 15 AA;

Query Match 87.1%; Score 61; DB 17; Length 15;  
 Best Local Similarity 83.3%; Pred. No. 0.0012;  
 Matches 10; Conservative 1; Mismatches 0; Gaps 0;

QY 1 CFAWKRNMRKVR 12  
 |||:|||||  
 Db 2 CFQWQRNMRKVR 13

Search completed: February 21, 2003, 07:56:44  
 Job time : 28.093 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:50:40 ; Search time 8.93023 Seconds  
(without alignments)  
39.537 Million cell updates/sec

Title: US-09-743-107B-94  
Perfect score: 70  
Sequence: 1 CFAWKRNMRKVR 12  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PTCUTS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	87.1	18	1	US-08-204-487-3
2	61	87.1	18	2	US-08-483-948-8
3	61	87.1	18	2	US-08-628-380-8
4	61	87.1	18	2	US-08-475-055-8
5	61	87.1	20	1	US-07-755-161A-3
6	61	87.1	20	1	US-07-891-174-3
7	61	87.1	20	1	US-08-204-487-1
8	61	87.1	20	1	US-08-256-771-24
9	61	87.1	20	1	US-08-256-771-25
10	61	87.1	20	1	US-08-381-984-24
11	61	87.1	20	1	US-08-381-984-25
12	61	87.1	22	4	US-09-508-734-4
13	61	87.1	24	4	US-09-508-734-6
14	61	87.1	25	1	US-07-755-161A-10
15	61	87.1	25	1	US-07-891-174-10
16	61	87.1	25	1	US-08-204-487-7
17	61	87.1	29	4	US-09-508-734-8
18	61	87.1	36	1	US-07-755-161A-8
19	61	87.1	36	1	US-07-891-174-8
20	61	87.1	36	1	US-08-256-771-30
21	61	87.1	36	1	US-08-381-984-29
22	61	87.1	47	2	US-08-484-182A-6
23	61	87.1	47	2	US-08-484-182A-7
24	61	87.1	50	2	US-08-693-274A-7
25	61	87.1	52	4	US-09-017-043A-3
26	61	87.1	53	2	US-08-464-182A-5
27	61	87.1	53	2	US-08-406-271-5

28	61	87.1	54	2	US-08-464-182A-2
29	61	87.1	54	2	US-08-406-271-2
30	61	87.1	594	3	US-08-724-586-2
31	61	87.1	694	4	US-09-421-632-2
32	61	87.1	694	4	US-09-932-190-2
33	61	87.1	705	2	US-08-655-640-2
34	61	87.1	708	2	US-08-655-640-4
35	61	87.1	711	1	US-08-154-019-4
36	61	87.1	711	1	US-08-461-333-4
37	61	87.1	711	3	US-08-464-167-4
38	61	87.1	711	3	US-08-158-313-4
39	61	87.1	711	4	US-08-476-798-4
40	58	82.9	711	1	US-08-145-681-2
41	58	82.9	711	1	US-08-250-308-2
42	58	82.9	711	1	US-08-453-703-2
43	58	82.9	711	2	US-08-456-106-2
44	58	82.9	711	3	US-08-456-108-2
45	58	82.9	711	4	US-09-265-577-2

ALIGNMENTS

RESULT 1  
US-08-204-487-3  
; Sequence 3, Application US/082044487  
; Patent No. 5565425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAOIKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGEKI  
; APPLICANT: DOSAKO, SHUN'ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIAKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; INHIBITORS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESSEE: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/08/204,487  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: RJN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..18  
; OTHER INFORMATION: /note= "HUMAN LACTOFERRIN PEPTIDE"

/ OTHER INFORMATION: (20-37)"  
US-08-204-487-3

Query Match 87.1%; Score 61; DB 1; Length 18;  
Best Local Similarity 83.3%; Pred. No. 0.00072;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CFAWKNNMKVR 12  
|||:|||||  
Db 1 CFQWQNNMKVR 12

## RESULT 2

US-08-485-948-8  
; Sequence 8, Application US/08485948  
; Patent No. 5855882  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,948  
; FILING DATE:  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/488,217  
; FILING DATE: JUNE 7, 1995  
; APPLICATION NUMBER: 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LF-C1, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
US-08-485-948-8

Query Match 87.1%; Score 61; DB 2; Length 18;  
Best Local Similarity 83.3%; Pred. No. 0.00072;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CFAWKNNMKVR 12  
|||:|||||  
Db 1 CFQWQNNMKVR 12

## RESULT 3

US-08-628-380-8  
; Sequence 8, Application US/08628380  
; Patent No. 5891341  
; GENERAL INFORMATION:  
; APPLICANT: LI, YONG MING  
; APPLICANT: VLASSARA, HELEN  
; APPLICANT: CERAMI, ANTHONY  
; TITLE OF INVENTION: AGE-MEDIATED INHIBITION OF ANTIBACTERIAL PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/628,380  
; FILING DATE: April 4, 1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LF-C1, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
US-08-628-380-8

Query Match 87.1%; Score 61; DB 2; Length 18;  
Best Local Similarity 83.3%; Pred. No. 0.00072;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CFAWKNNMKVR 12  
|||:|||||  
Db 1 CFQWQNNMKVR 12

## RESULT 4

US-08-475-055-8  
; Sequence 8, Application US/08475055  
; Patent No. 5962245  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Klauber & Jackson  
;; STREET: 411 Hackensack Avenue  
;; CITY: Hackensack  
;; STATE: New Jersey  
;; COUNTRY: USA  
;; ZIP: 07601  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; OPERATING SYSTEM: IBM PC compatible  
;; SOFTWARE: Patent in Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/475,055  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/485,948  
;; FILING DATE:  
;; APPLICATION NUMBER: 08/488,217  
;; FILING DATE: JUNE 7, 1995  
;; APPLICATION NUMBER: 08/418,642  
;; FILING DATE: APRIL 7, 1995  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Jackson Esq., David A.  
;; REGISTRATION NUMBER: 26,742  
;; REFERENCE/DOCKET NUMBER: 947-1-008A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 201 487-5800  
;; TELEFAX: 201 343-1684  
;; TELEX: 133521  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 18 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; DESCRIPTION: LF-Cl, 8-25  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FRAGMENT TYPE: internal  
US-08-475-055-8

Query Match 87.1%; Score 61; DB 2; Length 18;  
Best Local Similarity 83.3%; Pred. No. 0.00072;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFAWKRNMRKVR 12  
DB 1 CFWQNRNMRKVR 12

RESULT 5  
US-07-755-161A-3  
; Sequence 3, Application US/07755161A  
; Patent No. 5304633  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible

;; OPERATING SYSTEM: MS-DOS  
;; SOFTWARE: DisplayWrite  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/755,161A  
;; FILING DATE: 19910905  
;; CLASSIFICATION: 530  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Warren M. Cheek Jr.  
;; REGISTRATION NUMBER: 33,367  
;; REFERENCE/DOCKET NUMBER:  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-371-8850  
;; TELEFAX: 202-371-8856  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 20 amino acids  
;; TYPE: AMINO ACID  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE:  
;; HYPOTHETICAL:  
;; ANTI-SENSE:  
;; FRAGMENT TYPE:  
;; ORIGINAL SOURCE:  
;; ORGANISM:  
;; STRAIN:  
;; INDIVIDUAL ISOLATE:  
;; DEVELOPMENTAL STAGE:  
;; HAPLOTYPE:  
;; TISSUE TYPE:  
;; CELL TYPE:  
;; CELL LINE:  
;; ORGANELLE:  
;; IMMEDIATE SOURCE:  
;; CLONE:  
;; POSITION IN GENOME:  
;; CHROMOSOME/SEGMENT:  
;; MAP POSITION:  
;; UNITS:  
;; FEATURE:  
;; NAME/KEY: modified site  
;; LOCATION: 2  
;; IDENTIFICATION METHOD:  
;; OTHER INFORMATION: /note= "thiol group of  
;; OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with  
;; OTHER INFORMATION: thiol group of Cys residue at location 19"  
;; FEATURE:  
;; NAME/KEY: modified site  
;; LOCATION: 19  
;; IDENTIFICATION METHOD:  
;; OTHER INFORMATION: /note= "thiol group of  
;; OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with  
;; OTHER INFORMATION: thiol group of Cys residue at location 2"  
;; PUBLICATION INFORMATION:  
;; AUTHORS:  
;; TITLE:  
;; JOURNAL:  
;; VOLUME:  
;; ISSUE:  
;; DATE:  
;; DOCUMENT NUMBER:  
;; FILING DATE:  
;; PUBLICATION DATE:  
;; RELEVANT RESIDUES IN SEQ ID NO:  
US-07-755-161A-3

Query Match 87.1%; Score 61; DB 1; Length 20;

Best Local Similarity 83.3%; Pred. No. 0.00079;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFAWKNMKVR 12  
Db 2 CFQWQNNMKVR 13

RESULT 6  
US-07-891-174-3  
; Sequence 3, Application US/07891174  
; Patent No. 5317084  
; GENERAL INFORMATION:  
; APPLICANT: MATSUDA TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial Peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: DisplayWrite  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/891,174  
; FILING DATE: 29-MAY-1992  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/755,161  
; FILING DATE: 05-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX: 202-371-8856  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; HYPOTHETICAL:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:  
; CELL LINE:  
; ORGANELLER:  
; IMMEDIATE SOURCE:  
; LIBRARY:  
; CLONE:  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT:  
; MAP POSITION:  
; UNITS:  
; FEATURE:  
; NAME/KEY: modified site

LOCATION: 2  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 19"  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 19  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond wit  
OTHER INFORMATION: thiol group of Cys residue at location 2"  
PUBLICATION INFORMATION:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-891-174-3

Query Match 87.1%; Score 61; DB 1; Length 20;  
Best Local Similarity 83.3%; Pred. No. 0.00079;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFAWKNMKVR 12  
Db 2 CFQWQNNMKVR 13

RESULT 7  
US-08-204-487-1  
; Sequence 1, Application US/08204487  
; Patent No. 5565425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAOKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGERU  
; APPLICANT: DOSAKO, SHUN'ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; TITLE OF INVENTION: INHIBITORS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESSEE: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FJN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..20  
OTHER INFORMATION: /note= "ANTIBACTERIAL PEPTIDE  
DERIVED FROM HUMAN LACTOPERRIN"  
US-08-204-487-1

Query Match 87.1%; Score 61; DB 1; Length 20;  
Best Local Similarity 83.3%; Pred. No. 0.00079;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFAWKNNMKVR 12  
|||:|||||  
DB 2 CFQWQNNMKVR 13

RESULT 8  
US-08-256-771-24  
; Sequence 24, Application US/08256771  
; Patent No. 5656591  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
; TITLE OF INVENTION: PRODUCTS THEREWITH  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,771  
FILING DATE: July 22, 1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:

INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "Cys residues are linked by  
OTHER INFORMATION: disulfide bond"

US-08-256-771-24

Query Match 87.1%; Score 61; DB 1; Length 20;  
Best Local Similarity 83.3%; Pred. No. 0.00079;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFAWKNNMKVR 12  
|||:|||||  
DB 2 CFQWQNNMKVR 13

RESULT 9  
US-08-256-771-25  
; Sequence 25, Application US/08256771  
; Patent No. 5656591  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
; TITLE OF INVENTION: PRODUCTS THEREWITH  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,771  
FILING DATE: July 22, 1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "Cys residues are protected to  
OTHER INFORMATION: prevent disulfide bond"  
US-08-256-771-25

Query Match 87.1%; Score 61; DB 1; Length 20;  
Best Local Similarity 83.3%; Pred. No. 0.00079;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFAWKNNMKVR 12  
|||:|||||  
DB 2 CFQWQNNMKVR 13

RESULT 10  
US-08-381-984-24

Sequence 24, Application US/08381984

Patent No. 5804555  
 GENERAL INFORMATION:  
 APPLICANT: Mamoru TOMITA et al.  
 TITLE OF INVENTION: ANTIOXIDANT  
 NUMBER OF SEQUENCES: 32  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Wenderoth, Lind & Ponack  
 STREET: 805 Fifteenth Street, N.W., #700  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20005

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: MS-DOS  
 SOFTWARE: Wordperfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/381,984  
 FILING DATE: April 11, 1995  
 CLASSIFICATION: 252  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:

ATTORNEY/AGENT INFORMATION:  
 NAME: Warren M. Cheek, Jr.  
 REGISTRATION NUMBER: 33,367  
 REFERENCE/DOCKET NUMBER:  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-371-8850  
 TELEFAX:

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 20 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "cysteine residues at positions 2

OTHER INFORMATION: and 19 are bonded by disulfide linkage"

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "the specified peptide as well as

OTHER INFORMATION: peptides including the specified peptide as a fragment thereof

US-08-381-984-24

Query Match 87.1%; Score 61; DB 1; Length 20;

Best Local Similarity 83.3%; Pred. No. 0.00079;

Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFAWKRNMRKVR 12

DB 2 CFWQNRNRKVR 13

RESULT 11

US-08-381-984-25

Sequence 25, Application US/08381984

Patent No. 5804555  
 GENERAL INFORMATION:  
 APPLICANT: Mamoru TOMITA et al.  
 TITLE OF INVENTION: ANTIOXIDANT  
 NUMBER OF SEQUENCES: 32  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Wenderoth, Lind & Ponack

STREET: 805 Fifteenth Street, N.W., #700  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20005

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: MS-DOS  
 SOFTWARE: Wordperfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/381,984  
 FILING DATE: April 11, 1995  
 CLASSIFICATION: 252  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:

ATTORNEY/AGENT INFORMATION:  
 NAME: Warren M. Cheek, Jr.  
 REGISTRATION NUMBER: 33,367  
 REFERENCE/DOCKET NUMBER:  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-371-8850  
 TELEFAX:

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 20 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "the specified peptide as well as

OTHER INFORMATION: peptides including the specified peptide as a fragment the

FEATURE:

NAME/KEY:

LOCATION:

IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "cysteine residues at positions 2

OTHER INFORMATION: and 19 are chemically modified to prevent disulfide linkage

US-08-381-984-25

Query Match 87.1%; Score 61; DB 1; Length 20;

Best Local Similarity 83.3%; Pred. No. 0.00079;

Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFAWKRNMRKVR 12

DB 2 CFWQNRNRKVR 13

RESULT 12

US-09-508-734-4

Sequence 4, Application US/09508734

Patent No. 6423509

GENERAL INFORMATION:

APPLICANT: Samyang Genex Corporation

TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast a

TITLE OF INVENTION: useful microorganism thereof

FILE REFERENCE: PA/SYG/00139

CURRENT APPLICATION NUMBER: US/09/508,734

CURRENT FILING DATE: 2000-06-01

PRIOR FILING DATE: 1999-07-14

PRIOR APPLICATION NUMBER: FCI/KR99/00373

PRIOR FILING DATE: 1998-07-13

NUMBER OF SEQ ID NOS: 12

SOFTWARE: KopatentIn 1.71

SEQ ID NO 4

```
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-508-734-4

Query Match      87.1%; Score 61; DB 4; Length 22;
Best Local Similarity 83.3%; Pred. No. 0.00086;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFAWKNNRKVR 12
Db 2 CFWQWRNNRKVR 13

RESULT 13
US-09-508-734-6
; Sequence 6, Application US/09508734
; Patent No. 6423509
; GENERAL INFORMATION:
; APPLICANT: Samyang Genex Corporation
; TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and
; FILE REFERENCE: PA/SYG/00139
; CURRENT APPLICATION NUMBER: US/09/508,734
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/KR99/00373
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: KR1998-29351
; PRIOR FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Kopatentin 1.71
; SEQ ID NO 6
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-508-734-6

Query Match      87.1%; Score 61; DB 4; Length 24;
Best Local Similarity 83.3%; Pred. No. 0.00093;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFAWKNNRKVR 12
Db 3 CFWQWRNNRKVR 14

RESULT 14
US-07-755-161A-10
; Sequence 10, Application US/07755161A
; Patent No. 5304633
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; TITLE OF INVENTION: Antimicrobial Agent
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500KB
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/755,161A
; FILING DATE: 19910905
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHEetical:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; NAME/KEY: modified site
; LOCATION: 4
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 21"
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 21
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 4"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-07-755-161A-10

Query Match      87.1%; Score 61; DB 1; Length 25;
Best Local Similarity 83.3%; Pred. No. 0.00097;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFAWKNNRKVR 12
Db 4 CFWQWRNNRKVR 15
```



```
RESULT 15
US-07-891-174-10
; Sequence 10, Application US/07891174
; Patent No. 5317084
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; TITLE OF INVENTION: Antimicrobial Agent
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/891,174
; FILING DATE: 23-MAY-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/755,161
; FILING DATE: 05-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 4
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 21"
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 21
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 4"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-07-891-174-10
; Query Match 87.1%; Score 61; DB 1; Length 25;
; Best Local Similarity 83.3%; Pred. No. 0.00097;
; Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
; QY 1 CFWKKNMKVR 12
; DB 4 CFWKKNMKVR 15
; Search completed: February 21, 2003, 08:04:27
; Job time : 8.93023 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 08:08:15 ; Search time 6.88372 Seconds  
(without alignments)  
54.162 Million cell updates/sec

Title: US-09-743-107b-94

Perfect score: 70

Sequence: 1 CFAWKNRKR 12

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 156504 seqs, 31069816 residues

Total number of hits satisfying chosen parameters: 156504

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pdb:  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pdb:  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pdb:  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pdb:  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pdb:  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pdb:  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pdb:  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pdb:  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pdb:  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pdb:  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pdb:  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pdb:  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pdb:  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pdb:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	87.1	15	9	US-09-798-869-2
2	61	87.1	25	9	US-09-798-869-20
3	61	87.1	694	9	US-10-023-096-2
4	53	75.7	15	9	US-09-798-869-6
5	44	62.9	15	9	US-09-798-869-3
6	44	62.9	25	9	US-09-798-869-23
7	41	58.6	333	9	US-09-796-753-26
8	40	57.1	301	9	US-10-080-360-11
9	40	57.1	302	10	US-09-948-078-2
10	39	55.7	15	9	US-09-798-869-4
11	39	55.7	25	9	US-09-798-869-22
12	39	55.7	338	9	US-09-978-295A-119
13	39	55.7	338	9	US-09-978-697-119
14	39	55.7	338	9	US-09-978-192A-119
15	39	55.7	338	9	US-09-959-832A-119
16	39	55.7	338	9	US-09-978-189-119
17	39	55.7	338	9	US-09-976-753-14
18	39	55.7	553	10	US-09-981-649A-6
19	39	55.7	553	10	US-09-981-649A-24

20 55.7 554 10 US-09-981-649A-30 Sequence 30, Appl  
21 55.7 554 10 US-09-981-649A-32 Sequence 32, Appl  
22 55.7 559 10 US-09-981-649A-28 Sequence 28, Appl  
23 54.3 77 10 US-09-864-761-41002 Sequence 41002, A  
24 51.4 15 9 US-09-798-869-7 Sequence 7, Appl  
25 51.4 15 9 US-09-798-869-8 Sequence 8, Appl  
26 51.4 15 9 US-09-798-869-29 Sequence 29, Appl  
27 51.4 15 9 US-09-798-869-30 Sequence 30, Appl  
28 51.4 34 9 US-09-510-332-68 Sequence 68, Appl  
29 51.4 489 9 US-09-888-320-2 Sequence 2, Appl  
30 51.4 1013 9 US-10-028-072-38 Sequence 38, Appl  
31 51.4 1013 9 US-10-121-049-38 Sequence 38, Appl  
32 51.4 1013 9 US-10-123-904-38 Sequence 38, Appl  
33 51.4 1013 9 US-10-140-470-38 Sequence 38, Appl  
34 51.4 1013 9 US-10-175-746-38 Sequence 38, Appl  
35 51.4 1013 9 US-10-176-918-38 Sequence 38, Appl  
36 51.4 1013 9 US-10-176-921-38 Sequence 38, Appl  
37 51.4 1013 9 US-10-137-865-38 Sequence 38, Appl  
38 51.4 1013 9 US-10-140-474-38 Sequence 38, Appl  
39 50.0 46 10 US-09-864-761-35744 Sequence 35744, A  
40 50.0 184 10 US-09-925-301-1248 Sequence 1248, Ap  
41 48.6 81 10 US-09-925-300-1424 Sequence 1424, Ap  
42 48.6 86 9 US-09-738-626-5715 Sequence 5715, Ap  
43 48.6 302 10 US-09-915-582-50 Sequence 50, Appl  
44 48.6 509 10 US-09-879-957-194 Sequence 194, Appl  
45 48.6 1077 9 US-10-121-911-1 Sequence 1, Appl

#### ALIGNMENTS

RESULT 1  
US-09-798-869-2  
; Sequence 2, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version.4.0  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-09-798-869-2

Query Match 87.1%; Score 61; DB 9; Length 15;  
Best Local Similarity 83.3%; Pred. No. 0.00026;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFAWKNRKR 12  
|||:|||||  
Db 3 CFAWKNRKR 14

RESULT 2  
US-09-798-869-20  
; Sequence 20, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON

APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 25  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-798-869-20

Query Match 87.1%; Score 61; DB 9; Length 25;  
Best Local Similarity 83.3%; Pred. No. 0.00043;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQNNMKVR 12  
|||:|||||  
Db 3 CFQWQNNMKVR 14

RESULT 3  
US-10-023-096-2  
; Sequence 2, Application US/10023096  
; Patent No. US20020160941A1  
; GENERAL INFORMATION:  
; APPLICANT: Kruzel, Marian L.  
; APPLICANT: Kurecki, Tomasz  
; APPLICANT: Gollnick, Paul D.  
; APPLICANT: Doyle, Darrell J.  
; TITLE OF INVENTION: Cloning, Expression, and Uses of Human  
; TITLE OF INVENTION: Lactoferrin  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Jacobson, Price, Holman & Stern  
; STREET: 400 Seventh St. N.W.  
; CITY: Washington D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/023,096  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/724,586  
; FILING DATE: 30-SEPT-1996  
; APPLICATION NUMBER: US 08/238,445  
; FILING DATE: 05-MAY-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Player, William E.  
; REGISTRATION NUMBER: 31,409  
; REFERENCE/DOCKET NUMBER: 10505/P58185C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 638-6666  
; TELEFAX: (202) 393-5350  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 694 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-10-023-096-2

Query Match 87.1%; Score 61; DB 9; Length 694;  
Best Local Similarity 83.3%; Pred. No. 0.01; Indels 0; Gaps 0;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQWQNNMKVR 12  
|||:|||||  
Db 22 CFQWQNNMKVR 33

RESULT 4  
US-09-798-869-6  
; Sequence 6, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINER (RNSON  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of Homo sapiens  
; OTHER INFORMATION: sequence)  
US-09-798-869-6

Query Match 75.7%; Score 53; DB 9; Length 15;  
Best Local Similarity 75.0%; Pred. No. 0.0056;  
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQNNMKVR 12  
|||:|||||  
Db 3 CFQWQNNMKVR 14

RESULT 5  
US-09-798-869-3  
; Sequence 3, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINER (RNSON  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: CAPRINE  
US-09-798-869-3

Query Match 62.9%; Score 44; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.18;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFAWKNNRKY 11  
|:|:|:|:  
Db 3 CYQWQRMEKL 13

## RESULT 6

US-09-798-869-23  
; Sequence 23, Application US/09798869  
; Publication No. US2003002821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVINBJ (RMSO)  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; PRIOR FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/G899/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 25  
; TIPS: PRT  
; ORGANISM: CAPRINE  
US-09-798-869-23

Query Match 62.9%; Score 44; DB 9; Length 25;  
Best Local Similarity 54.5%; Pred. No. 0.28;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFAWKNNRKY 11  
|:|:|:|:  
Db 3 CYQWQRMEKL 13

## RESULT 7

US-09-796-753-26  
; Sequence 26, Application US/09796753  
; Publication No. US20030027998A1  
; GENERAL INFORMATION:  
; APPLICANT: McCarthy, Sean A.  
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF  
; FILE REFERENCE: 7853-227-999  
; CURRENT APPLICATION NUMBER: US/09/796,753  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 09/183,175  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: 09/223,094  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/223,546  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/224,246  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/259,388  
; PRIOR FILING DATE: 1999-02-26  
; PRIOR APPLICATION NUMBER: 60/122,458  
; PRIOR FILING DATE: 1999-03-01  
; PRIOR APPLICATION NUMBER: 09/312,359  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 09/336,536  
; PRIOR FILING DATE: 1999-06-18  
; PRIOR APPLICATION NUMBER: 09/342,687  
; PRIOR FILING DATE: 1999-06-29  
; PRIOR APPLICATION NUMBER: 09/345,464  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: 09/365,164  
; PRIOR FILING DATE: 1999-07-30

; PRIOR APPLICATION NUMBER: 09/399,723  
; PRIOR FILING DATE: 1999-09-20  
; PRIOR APPLICATION NUMBER: 09/409,634  
; PRIOR FILING DATE: 1999-09-30  
; PRIOR APPLICATION NUMBER: 09/471,179  
; PRIOR FILING DATE: 1999-12-23  
; PRIOR APPLICATION NUMBER: 09/474,071  
; PRIOR FILING DATE: 1999-12-29  
; PRIOR APPLICATION NUMBER: 09/474,072  
; PRIOR FILING DATE: 1999-12-29  
; PRIOR APPLICATION NUMBER: 09/514,010  
; PRIOR FILING DATE: 2000-02-25  
; PRIOR APPLICATION NUMBER: 09/516,745  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 09/572,002  
; PRIOR FILING DATE: 2000-05-14  
; PRIOR APPLICATION NUMBER: 09/597,993  
; PRIOR FILING DATE: 2000-06-19  
; PRIOR APPLICATION NUMBER: 09/599,596  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 09/630,334  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: 09/606,565  
; PRIOR FILING DATE: 2000-06-29  
; PRIOR APPLICATION NUMBER: 09/606,317  
; PRIOR FILING DATE: 2000-06-29  
; PRIOR APPLICATION NUMBER: 09/665,666  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: 09/677,751  
; PRIOR FILING DATE: 2000-09-30  
; NUMBER OF SEQ ID NOS: 162  
; SEQ ID NO 26  
; LENGTH: 333  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-796-753-26

Query Match 58.6%; Score 41; DB 9; Length 333;  
Best Local Similarity 54.5%; Pred. No. 10;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFAWKNNRKY 11  
|:|:|:|:  
Db 48 CYGWRNNKGV 58

## RESULT 8

US-10-080-960-11  
; Sequence 11, Application US/10080960  
; Publication No. US20020197695A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Glucksmann, Maria  
; APPLICANT: Meyers, Rachel  
; TITLE OF INVENTION: 80090, 52874, 52880, 63497, AND 33425  
; FILE REFERENCE: 38155-20044.00  
; CURRENT APPLICATION NUMBER: US/10/080,960  
; CURRENT FILING DATE: 2001-10-19  
; PRIOR APPLICATION NUMBER: US 60/242,040  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: US 60/242,038  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: US 60/241,992  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: US 60/242,637  
; PRIOR FILING DATE: 2000-10-23  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 301  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-10-080-960-11

Query Match 57.1%; Score 40; DB 9; Length 301;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFAWRK 6  
 |||||  
 DB 296 CFAWRK 301

RESULT 9

US-09-948-078-2  
 ; Sequence 2, Application US/09948078  
 ; Patent No. US20020147308A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HOLLOWAY, JAMES L.  
 ; APPLICANT: LOK, SI  
 ; TITLE OF INVENTION: Human Vomeronasal Receptor  
 ; FILE REFERENCE: 00-73  
 ; CURRENT APPLICATION NUMBER: US/09/948,078  
 ; CURRENT FILING DATE: 2001-09-06  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 302  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-948-078-2

Query Match 57.1%; Score 40; DB 10; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFAWRK 6  
 |||||  
 DB 297 CFAWRK 302

RESULT 10

US-09-798-869-4  
 ; Sequence 4, Application US/09798869  
 ; Publication No. US2003002821A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOHN SIGURD SVENDSEN  
 ; APPLICANT: (YSTEIN REKDAL  
 ; APPLICANT: BALDUR SVEINBJ(RNSSON  
 ; APPLICANT: LARS VORLAND  
 ; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
 ; FILE REFERENCE: A34049-PCT-USA-A  
 ; CURRENT APPLICATION NUMBER: US/09/798,869  
 ; CURRENT FILING DATE: 2001-02-27  
 ; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
 ; PRIOR FILING DATE: 1999-08-31  
 ; PRIOR APPLICATION NUMBER: GB9818938.4  
 ; PRIOR FILING DATE: 1998-08-28  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 4  
 ; LENGTH: 15  
 ; TYPE: PRT  
 ; ORGANISM: MURINE  
 US-09-798-869-4

Query Match 55.7%; Score 39; DB 9; Length 15;  
 Best Local Similarity 54.5%; Pred. No. 1.2;  
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFAWRK 11  
 |||||  
 DB 3 CLRQWENRKY 13

RESULT 11

US-09-798-869-22  
 ; Sequence 22, Application US/09798869  
 ; Publication No. US2003002821A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOHN SIGURD SVENDSEN  
 ; APPLICANT: (YSTEIN REKDAL  
 ; APPLICANT: BALDUR SVEINBJ(RNSSON  
 ; APPLICANT: LARS VORLAND  
 ; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
 ; FILE REFERENCE: A34049-PCT-USA-A  
 ; CURRENT APPLICATION NUMBER: US/09/798,869  
 ; CURRENT FILING DATE: 2001-02-27  
 ; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
 ; PRIOR FILING DATE: 1999-08-31  
 ; PRIOR APPLICATION NUMBER: GB9818938.4  
 ; PRIOR FILING DATE: 1998-08-28  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 22  
 ; LENGTH: 25  
 ; TYPE: PRT  
 ; ORGANISM: MURINE  
 US-09-798-869-22

Query Match 55.7%; Score 39; DB 9; Length 25;  
 Best Local Similarity 54.5%; Pred. No. 1.9;  
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFAWRK 11  
 |||||  
 DB 3 CLRQWENRKY 13

RESULT 12

US-09-978-295A-119  
 ; Sequence 119, Application US/09978295A  
 ; Patent No. US20020156006A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi  
 ; APPLICANT: Baker Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnovers, Luc  
 ; APPLICANT: Eaton, Dan  
 ; APPLICANT: Ferrara, Napoleon  
 ; APPLICANT: Filvarsoff, Ellen  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Hillan, Kenneth J.  
 ; APPLICANT: Kijavini, Ivar J.  
 ; APPLICANT: Kuo, Sophia S.  
 ; APPLICANT: Napier, Mary A.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Paoni, Nicholas F.  
 ; APPLICANT: Roy, Margaret Ann  
 ; APPLICANT: Shelton, David L.  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Williams, P. Mickey  
 ; APPLICANT: Wood, William I.  
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 ; FILE REFERENCE: P2630PIC11  
 ; CURRENT APPLICATION NUMBER: US/09/978,295A  
 ; CURRENT FILING DATE: 2001-10-15  
 ; PRIOR APPLICATION NUMBER: 09/918585  
 ; PRIOR FILING DATE: 2001-07-30

1 PRIOR APPLICATION NUMBER: 60/062250  
2 PRIOR FILING DATE: 1997-10-17  
3 PRIOR APPLICATION NUMBER: 60/064249  
4 PRIOR FILING DATE: 1997-11-03  
5 PRIOR APPLICATION NUMBER: 60/065311  
6 PRIOR FILING DATE: 1997-11-13  
7 PRIOR APPLICATION NUMBER: 60/066364  
8 PRIOR FILING DATE: 1997-11-21  
9 PRIOR APPLICATION NUMBER: 60/077450  
10 PRIOR FILING DATE: 1998-03-10  
11 PRIOR APPLICATION NUMBER: 60/077632  
12 PRIOR FILING DATE: 1998-03-11  
13 PRIOR APPLICATION NUMBER: 60/077641  
14 PRIOR FILING DATE: 1998-03-11  
15 PRIOR APPLICATION NUMBER: 60/077649  
16 PRIOR FILING DATE: 1998-03-11  
17 PRIOR APPLICATION NUMBER: 60/077791  
18 PRIOR FILING DATE: 1998-03-12  
19 PRIOR APPLICATION NUMBER: 60/078004  
20 PRIOR FILING DATE: 1998-03-13  
21 PRIOR APPLICATION NUMBER: 60/078886  
22 PRIOR FILING DATE: 1998-03-20  
23 PRIOR APPLICATION NUMBER: 60/078936  
24 PRIOR FILING DATE: 1998-03-20  
25 PRIOR APPLICATION NUMBER: 60/078910  
26 PRIOR FILING DATE: 1998-03-20  
27 PRIOR APPLICATION NUMBER: 60/078939  
28 PRIOR FILING DATE: 1998-03-20  
29 PRIOR APPLICATION NUMBER: 60/079294  
30 PRIOR FILING DATE: 1998-03-25  
31 PRIOR APPLICATION NUMBER: 60/079656  
32 PRIOR FILING DATE: 1998-03-26  
33 PRIOR APPLICATION NUMBER: 60/079664  
34 PRIOR FILING DATE: 1998-03-27  
35 PRIOR APPLICATION NUMBER: 60/079689  
36 PRIOR FILING DATE: 1998-03-27  
37 PRIOR APPLICATION NUMBER: 60/079663  
38 PRIOR FILING DATE: 1998-03-27  
39 PRIOR APPLICATION NUMBER: 60/079728  
40 PRIOR FILING DATE: 1998-03-27  
41 PRIOR APPLICATION NUMBER: 60/079786  
42 PRIOR FILING DATE: 1998-03-27  
43 PRIOR APPLICATION NUMBER: 60/079920  
44 PRIOR FILING DATE: 1998-03-30  
45 PRIOR APPLICATION NUMBER: 60/079923  
46 PRIOR FILING DATE: 1998-03-30  
47 PRIOR APPLICATION NUMBER: 60/080105  
48 PRIOR FILING DATE: 1998-03-31  
49 PRIOR APPLICATION NUMBER: 60/080107  
50 PRIOR FILING DATE: 1998-03-31  
51 PRIOR APPLICATION NUMBER: 60/080165  
52 PRIOR FILING DATE: 1998-03-31  
53 PRIOR APPLICATION NUMBER: 60/080194  
54 PRIOR FILING DATE: 1998-03-31  
55 PRIOR APPLICATION NUMBER: 60/080327  
56 PRIOR FILING DATE: 1998-04-01  
57 PRIOR APPLICATION NUMBER: 60/080328  
58 PRIOR FILING DATE: 1998-04-01  
59 PRIOR APPLICATION NUMBER: 60/080333  
60 PRIOR FILING DATE: 1998-04-01  
61 PRIOR APPLICATION NUMBER: 60/080334  
62 PRIOR FILING DATE: 1998-04-01  
63 PRIOR APPLICATION NUMBER: 60/081070  
64 PRIOR FILING DATE: 1998-04-08  
65 PRIOR APPLICATION NUMBER: 60/081049  
66 PRIOR FILING DATE: 1998-04-08  
67 PRIOR APPLICATION NUMBER: 60/081071  
68 PRIOR FILING DATE: 1998-04-08  
69 PRIOR APPLICATION NUMBER: 60/081195  
70 PRIOR FILING DATE: 1998-04-08  
71 PRIOR APPLICATION NUMBER: 60/081203  
72 PRIOR FILING DATE: 1998-04-09  
73 PRIOR APPLICATION NUMBER: 60/081229  
74 PRIOR FILING DATE: 1998-04-09  
75 PRIOR APPLICATION NUMBER: 60/081955  
76 PRIOR FILING DATE: 1998-04-15  
77 PRIOR APPLICATION NUMBER: 60/081817  
78 PRIOR FILING DATE: 1998-04-15  
79 PRIOR APPLICATION NUMBER: 60/081819  
80 PRIOR FILING DATE: 1998-04-15  
81 PRIOR APPLICATION NUMBER: 60/081952  
82 PRIOR FILING DATE: 1998-04-15  
83 PRIOR APPLICATION NUMBER: 60/081838  
84 PRIOR FILING DATE: 1998-04-15  
85 PRIOR APPLICATION NUMBER: 60/082568  
86 PRIOR FILING DATE: 1998-04-21  
87 PRIOR APPLICATION NUMBER: 60/082569  
88 PRIOR FILING DATE: 1998-04-21  
89 PRIOR APPLICATION NUMBER: 60/082704  
90 PRIOR FILING DATE: 1998-04-22  
91 PRIOR APPLICATION NUMBER: 60/082804  
92 PRIOR FILING DATE: 1998-04-22  
93 PRIOR APPLICATION NUMBER: 60/082700  
94 PRIOR FILING DATE: 1998-04-22  
95 PRIOR APPLICATION NUMBER: 60/082797  
96 PRIOR FILING DATE: 1998-04-22  
97 PRIOR APPLICATION NUMBER: 60/082796  
98 PRIOR FILING DATE: 1998-04-23  
99 PRIOR APPLICATION NUMBER: 60/083336  
100 PRIOR FILING DATE: 1998-04-27  
101 PRIOR APPLICATION NUMBER: 60/083322  
102 PRIOR FILING DATE: 1998-04-28  
103 PRIOR APPLICATION NUMBER: 60/083392  
104 PRIOR FILING DATE: 1998-04-29  
105 PRIOR APPLICATION NUMBER: 60/083495  
106 PRIOR FILING DATE: 1998-04-29  
107 PRIOR APPLICATION NUMBER: 60/083496  
108 PRIOR FILING DATE: 1998-04-29  
109 PRIOR APPLICATION NUMBER: 60/083499  
110 PRIOR FILING DATE: 1998-04-29  
111 PRIOR APPLICATION NUMBER: 60/083545  
112 PRIOR FILING DATE: 1998-04-29  
113 PRIOR APPLICATION NUMBER: 60/083554  
114 PRIOR FILING DATE: 1998-04-29  
115 PRIOR APPLICATION NUMBER: 60/083558  
116 PRIOR FILING DATE: 1998-04-29  
117 PRIOR APPLICATION NUMBER: 60/083559  
118 PRIOR FILING DATE: 1998-04-29  
119 PRIOR APPLICATION NUMBER: 60/083500  
120 PRIOR FILING DATE: 1998-04-29  
121 PRIOR APPLICATION NUMBER: 60/083742  
122 PRIOR FILING DATE: 1998-04-30  
123 PRIOR APPLICATION NUMBER: 60/084366  
124 PRIOR FILING DATE: 1998-05-05  
125 PRIOR APPLICATION NUMBER: 60/084414  
126 PRIOR FILING DATE: 1998-05-06  
127 PRIOR APPLICATION NUMBER: 60/084441  
128 PRIOR FILING DATE: 1998-05-06  
129 PRIOR APPLICATION NUMBER: 60/084637  
130 PRIOR FILING DATE: 1998-05-07  
131 PRIOR APPLICATION NUMBER: 60/084639  
132 PRIOR FILING DATE: 1998-05-07  
133 PRIOR APPLICATION NUMBER: 60/084640  
134 PRIOR FILING DATE: 1998-05-07  
135 PRIOR APPLICATION NUMBER: 60/084598  
136 PRIOR FILING DATE: 1998-05-07  
137 PRIOR APPLICATION NUMBER: 60/084600  
138 PRIOR FILING DATE: 1998-05-07  
139 PRIOR APPLICATION NUMBER: 60/084627  
140 PRIOR FILING DATE: 1998-05-07  
141 PRIOR APPLICATION NUMBER: 60/084643  
142 PRIOR FILING DATE: 1998-05-07  
143 PRIOR APPLICATION NUMBER: 60/085339  
144 PRIOR FILING DATE: 1998-05-13  
145 PRIOR APPLICATION NUMBER: 60/085338  
146 PRIOR FILING DATE: 1998-05-13

; PRIOR APPLICATION NUMBER: 60/085323  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085582  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085700  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085689  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085579  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085580  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085573  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085704  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085697

Query Match 55.7%; Score 39; DB 9; Length 338;  
 Best Local Similarity 45.5%; Pred. No. 23;  
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFAWRNRKV 11  
 Db 50 CYGWRNRSGV 60

## RESULT 13

; Sequence 119, Application US/09978697  
 ; Patent No. US20020169284A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi  
 ; APPLICANT: Baker Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnovers, Luc  
 ; APPLICANT: Eaton, Dan  
 ; APPLICANT: Ferrara, Napoleon  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Hillan, Kenneth J.  
 ; APPLICANT: Kljavin, Ivar J.  
 ; APPLICANT: Kuo, Sophia S.  
 ; APPLICANT: Napier, Mary A.  
 ; APPLICANT: Pan, James;  
 ; APPLICANT: Paoni, Nicholas F.  
 ; APPLICANT: Roy, Margaret Ann  
 ; APPLICANT: Shelton, David L.  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tamas, Daniel  
 ; APPLICANT: Williams, P. Mickey  
 ; APPLICANT: Wood, William I.  
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 ; TITLE OF INVENTION: Acids Encoding the Same  
 ; FILE REFERENCE: P2630P1C27  
 ; CURRENT APPLICATION NUMBER: US/09/978,697  
 ; CURRENT FILING DATE: 2001-10-16  
 ; PRIOR APPLICATION NUMBER: 09/918585  
 ; PRIOR FILING DATE: 2001-07-30  
 ; PRIOR APPLICATION NUMBER: 60/062250  
 ; PRIOR FILING DATE: 1997-10-17  
 ; PRIOR APPLICATION NUMBER: 60/064249  
 ; PRIOR FILING DATE: 1997-11-03  
 ; PRIOR APPLICATION NUMBER: 60/065311  
 ; PRIOR FILING DATE: 1997-11-13  
 ; PRIOR APPLICATION NUMBER: 60/066364

; PRIOR FILING DATE: 1997-11-21  
 ; PRIOR APPLICATION NUMBER: 60/077450  
 ; PRIOR FILING DATE: 1998-03-10  
 ; PRIOR APPLICATION NUMBER: 60/077632  
 ; PRIOR FILING DATE: 1998-03-11  
 ; PRIOR APPLICATION NUMBER: 60/077641  
 ; PRIOR FILING DATE: 1998-03-11  
 ; PRIOR APPLICATION NUMBER: 60/077649  
 ; PRIOR FILING DATE: 1998-03-11  
 ; PRIOR APPLICATION NUMBER: 60/077791  
 ; PRIOR FILING DATE: 1998-03-12  
 ; PRIOR APPLICATION NUMBER: 60/078004  
 ; PRIOR FILING DATE: 1998-03-13  
 ; PRIOR APPLICATION NUMBER: 60/078886  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/078936  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/078910  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/078939  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/079294  
 ; PRIOR FILING DATE: 1998-03-25  
 ; PRIOR APPLICATION NUMBER: 60/079656  
 ; PRIOR FILING DATE: 1998-03-26  
 ; PRIOR APPLICATION NUMBER: 60/079664  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079689  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079663  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079728  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079786  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079920  
 ; PRIOR FILING DATE: 1998-03-30  
 ; PRIOR APPLICATION NUMBER: 60/079923  
 ; PRIOR FILING DATE: 1998-03-30  
 ; PRIOR APPLICATION NUMBER: 60/080105  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080107  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080165  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080194  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080327  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/080328  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/080333  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/080334  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/081070  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081049  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081071  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081195  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081203  
 ; PRIOR FILING DATE: 1998-04-09  
 ; PRIOR APPLICATION NUMBER: 60/081229  
 ; PRIOR FILING DATE: 1998-04-09  
 ; PRIOR APPLICATION NUMBER: 60/081955  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081817  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081819  
 ; PRIOR FILING DATE: 1998-04-15

;; PRIOR APPLICATION NUMBER: 60/081952  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081838  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/082568  
;; PRIOR FILING DATE: 1998-04-21  
;; PRIOR APPLICATION NUMBER: 60/082569  
;; PRIOR FILING DATE: 1998-04-21  
;; PRIOR APPLICATION NUMBER: 60/082704  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082804  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082700  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082797  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082796  
;; PRIOR FILING DATE: 1998-04-23  
;; PRIOR APPLICATION NUMBER: 60/083336  
;; PRIOR FILING DATE: 1998-04-27  
;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/083392  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083495  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083496  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083499  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083545  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083554  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083558  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083559  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083500  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083742  
;; PRIOR FILING DATE: 1998-04-30  
;; PRIOR APPLICATION NUMBER: 60/084366  
;; PRIOR FILING DATE: 1998-05-05  
;; PRIOR APPLICATION NUMBER: 60/084414  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084441  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084637  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084639  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084640  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084598  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084643  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/083339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/083338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085582  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085700  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085689

;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085580  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085573  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697

Quary Match 55.7%; Score 39; DB 9; Length 338;  
Best Local Similarity 45.5%; Pred.No. 23;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFAKRNMRKV 11  
| : | : | : |  
Db 50 CYGWRNSKGV 60

## RESULT 14

US-09-978-192A-119  
; Sequence 119, Application US/09978192A  
; Patent No. US2002017753A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Baker Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan  
; APPLICANT: Ferrara, Napoleon  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Kuo, Sophia S.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James;  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Shelton, David L.  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2630P1C9  
; CURRENT APPLICATION NUMBER: US/09/978,192A  
; CURRENT FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: 09/918585  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/064249  
; PRIOR FILING DATE: 1997-11-03  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066364  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: 60/077450  
; PRIOR FILING DATE: 1998-03-10  
; PRIOR APPLICATION NUMBER: 60/077632  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077641  
; PRIOR FILING DATE: 1998-03-11



1 PRIOR APPLICATION NUMBER: 60/077649  
2 PRIOR FILING DATE: 1998-03-11  
3 PRIOR APPLICATION NUMBER: 60/077791  
4 PRIOR FILING DATE: 1998-03-12  
5 PRIOR APPLICATION NUMBER: 60/078004  
6 PRIOR FILING DATE: 1998-03-13  
7 PRIOR APPLICATION NUMBER: 60/078886  
8 PRIOR FILING DATE: 1998-03-20  
9 PRIOR APPLICATION NUMBER: 60/078936  
10 PRIOR FILING DATE: 1998-03-20  
11 PRIOR APPLICATION NUMBER: 60/078910  
12 PRIOR FILING DATE: 1998-03-20  
13 PRIOR APPLICATION NUMBER: 60/078939  
14 PRIOR FILING DATE: 1998-03-20  
15 PRIOR APPLICATION NUMBER: 60/079294  
16 PRIOR FILING DATE: 1998-03-25  
17 PRIOR APPLICATION NUMBER: 60/079656  
18 PRIOR FILING DATE: 1998-03-26  
19 PRIOR APPLICATION NUMBER: 60/079664  
20 PRIOR FILING DATE: 1998-03-27  
21 PRIOR APPLICATION NUMBER: 60/079699  
22 PRIOR FILING DATE: 1998-03-27  
23 PRIOR APPLICATION NUMBER: 60/079663  
24 PRIOR FILING DATE: 1998-03-27  
25 PRIOR APPLICATION NUMBER: 60/079728  
26 PRIOR FILING DATE: 1998-03-27  
27 PRIOR APPLICATION NUMBER: 60/079786  
28 PRIOR FILING DATE: 1998-03-27  
29 PRIOR APPLICATION NUMBER: 60/079920  
30 PRIOR FILING DATE: 1998-03-30  
31 PRIOR APPLICATION NUMBER: 60/079923  
32 PRIOR FILING DATE: 1998-03-30  
33 PRIOR APPLICATION NUMBER: 60/080105  
34 PRIOR FILING DATE: 1998-03-31  
35 PRIOR APPLICATION NUMBER: 60/080107  
36 PRIOR FILING DATE: 1998-03-31  
37 PRIOR APPLICATION NUMBER: 60/080165  
38 PRIOR FILING DATE: 1998-03-31  
39 PRIOR APPLICATION NUMBER: 60/080194  
40 PRIOR FILING DATE: 1998-03-31  
41 PRIOR APPLICATION NUMBER: 60/080327  
42 PRIOR FILING DATE: 1998-04-01  
43 PRIOR APPLICATION NUMBER: 60/080328  
44 PRIOR FILING DATE: 1998-04-01  
45 PRIOR APPLICATION NUMBER: 60/080333  
46 PRIOR FILING DATE: 1998-04-01  
47 PRIOR APPLICATION NUMBER: 60/080334  
48 PRIOR FILING DATE: 1998-04-01  
49 PRIOR APPLICATION NUMBER: 60/081070  
50 PRIOR FILING DATE: 1998-04-08  
51 PRIOR APPLICATION NUMBER: 60/081049  
52 PRIOR FILING DATE: 1998-04-08  
53 PRIOR APPLICATION NUMBER: 60/081071  
54 PRIOR FILING DATE: 1998-04-08  
55 PRIOR APPLICATION NUMBER: 60/081195  
56 PRIOR FILING DATE: 1998-04-08  
57 PRIOR APPLICATION NUMBER: 60/081203  
58 PRIOR FILING DATE: 1998-04-09  
59 PRIOR APPLICATION NUMBER: 60/081229  
60 PRIOR FILING DATE: 1998-04-09  
61 PRIOR APPLICATION NUMBER: 60/081955  
62 PRIOR FILING DATE: 1998-04-15  
63 PRIOR APPLICATION NUMBER: 60/081817  
64 PRIOR FILING DATE: 1998-04-15  
65 PRIOR APPLICATION NUMBER: 60/081819  
66 PRIOR FILING DATE: 1998-04-15  
67 PRIOR APPLICATION NUMBER: 60/081952  
68 PRIOR FILING DATE: 1998-04-15  
69 PRIOR APPLICATION NUMBER: 60/081838  
70 PRIOR FILING DATE: 1998-04-15  
71 PRIOR APPLICATION NUMBER: 60/082568  
72 PRIOR FILING DATE: 1998-04-21  
73 PRIOR APPLICATION NUMBER: 60/082569

74 PRIOR FILING DATE: 1998-04-21  
75 PRIOR APPLICATION NUMBER: 60/082704  
76 PRIOR FILING DATE: 1998-04-22  
77 PRIOR APPLICATION NUMBER: 60/082804  
78 PRIOR FILING DATE: 1998-04-22  
79 PRIOR APPLICATION NUMBER: 60/082700  
80 PRIOR FILING DATE: 1998-04-22  
81 PRIOR APPLICATION NUMBER: 60/082797  
82 PRIOR FILING DATE: 1998-04-22  
83 PRIOR APPLICATION NUMBER: 60/082796  
84 PRIOR FILING DATE: 1998-04-23  
85 PRIOR APPLICATION NUMBER: 60/083336  
86 PRIOR FILING DATE: 1998-04-27  
87 PRIOR APPLICATION NUMBER: 60/083322  
88 PRIOR FILING DATE: 1998-04-28  
89 PRIOR APPLICATION NUMBER: 60/083392  
90 PRIOR FILING DATE: 1998-04-29  
91 PRIOR APPLICATION NUMBER: 60/083495  
92 PRIOR FILING DATE: 1998-04-29  
93 PRIOR APPLICATION NUMBER: 60/083496  
94 PRIOR FILING DATE: 1998-04-29  
95 PRIOR APPLICATION NUMBER: 60/083499  
96 PRIOR FILING DATE: 1998-04-29  
97 PRIOR APPLICATION NUMBER: 60/083545  
98 PRIOR FILING DATE: 1998-04-29  
99 PRIOR APPLICATION NUMBER: 60/083554  
100 PRIOR FILING DATE: 1998-04-29  
101 PRIOR APPLICATION NUMBER: 60/083558  
102 PRIOR FILING DATE: 1998-04-29  
103 PRIOR APPLICATION NUMBER: 60/083559  
104 PRIOR FILING DATE: 1998-04-29  
105 PRIOR APPLICATION NUMBER: 60/083500  
106 PRIOR FILING DATE: 1998-04-29  
107 PRIOR APPLICATION NUMBER: 60/083742  
108 PRIOR FILING DATE: 1998-04-30  
109 PRIOR APPLICATION NUMBER: 60/084366  
110 PRIOR FILING DATE: 1998-05-05  
111 PRIOR APPLICATION NUMBER: 60/084414  
112 PRIOR FILING DATE: 1998-05-06  
113 PRIOR APPLICATION NUMBER: 60/084441  
114 PRIOR FILING DATE: 1998-05-06  
115 PRIOR APPLICATION NUMBER: 60/084637  
116 PRIOR FILING DATE: 1998-05-07  
117 PRIOR APPLICATION NUMBER: 60/084639  
118 PRIOR FILING DATE: 1998-05-07  
119 PRIOR APPLICATION NUMBER: 60/084640  
120 PRIOR FILING DATE: 1998-05-07  
121 PRIOR APPLICATION NUMBER: 60/084598  
122 PRIOR FILING DATE: 1998-05-07  
123 PRIOR APPLICATION NUMBER: 60/084600  
124 PRIOR FILING DATE: 1998-05-07  
125 PRIOR APPLICATION NUMBER: 60/084627  
126 PRIOR FILING DATE: 1998-05-07  
127 PRIOR APPLICATION NUMBER: 60/084643  
128 PRIOR FILING DATE: 1998-05-07  
129 PRIOR APPLICATION NUMBER: 60/085339  
130 PRIOR FILING DATE: 1998-05-13  
131 PRIOR APPLICATION NUMBER: 60/085338  
132 PRIOR FILING DATE: 1998-05-13  
133 PRIOR APPLICATION NUMBER: 60/085323  
134 PRIOR FILING DATE: 1998-05-13  
135 PRIOR APPLICATION NUMBER: 60/085582  
136 PRIOR FILING DATE: 1998-05-15  
137 PRIOR APPLICATION NUMBER: 60/085700  
138 PRIOR FILING DATE: 1998-05-15  
139 PRIOR APPLICATION NUMBER: 60/085689  
140 PRIOR FILING DATE: 1998-05-15  
141 PRIOR APPLICATION NUMBER: 60/085579  
142 PRIOR FILING DATE: 1998-05-15  
143 PRIOR APPLICATION NUMBER: 60/085580  
144 PRIOR FILING DATE: 1998-05-15  
145 PRIOR APPLICATION NUMBER: 60/085573  
146 PRIOR FILING DATE: 1998-05-15

```
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match      55.7%; Score 39; DB 9; Length 338;
Best Local Similarity 45.5%; Pred. No. 23;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 CFAWKNNRKY 11
Db 50 CYGWRNRSGV 60

RESULT 15
US-09-999-832A-119
; Sequence 119, Application US/09999832A
; Publication No. US20020192706A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tunas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P263021C63
; CURRENT APPLICATION NUMBER: US/09/999,832A
; CURRENT FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
```

1 CFAKRNMRKV 11  
: : : : :  
Db 50 CYGWRNSKGV 60

Search completed: February 21, 2003, 08:11:57  
Job time : 6.88372 secs

PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084639  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084640  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084598  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084643  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085582  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085700  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085689  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085580  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085573  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085704  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697

Query Match 55.7%; Score 39; DB 9; Length 338;  
Best Local Similarity 45.5%; Pred. No. 23;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:48:01 ; Search time 10.6047 Seconds  
(without alignments)  
108.784 Million cell updates/sec

Title: US-09-743-107B-94  
Perfect score: 70  
Sequence: 1 CFAWKRNMRKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 73.\*

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	87.1	711	1 THUL	lactotransferrin p
2	46	65.7	511	2 AB0858	hypothetical prote
3	44	62.9	708	2 J02323	lactoferrin - goat
4	43	61.4	393	2 B64639	3-deoxy-manno-octu
5	41	58.6	33	2 S52107	lactoferrin - shee
6	41	58.6	584	2 C84325	hypothetical prote
7	40	57.1	267	2 S78822	hypothetical prote
8	40	57.1	386	2 B82921	serine/threonine k
9	39	55.7	114	2 D33876	carcinoembryonic a
10	39	55.7	121	2 AH3147	hypothetical prote
11	39	55.7	206	2 H97451	pyridoxamine 5'-ph
12	39	55.7	206	2 A82670	probable pyridoxam
13	39	55.7	208	2 AG3441	hypothetical prote
14	39	55.7	275	2 T22597	hypothetical prote
15	39	55.7	282	2 F90580	hypothetical prote
16	39	55.7	393	2 D71876	3-deoxy-manno-octu
17	39	55.7	447	2 T18633	hypothetical prote
18	39	55.7	558	2 T17324	hypothetical prote
19	39	55.7	675	2 S51037	zinc-finger protei
20	39	55.7	707	1 A28438	lactoferrin precu
21	39	55.7	720	2 A36807	hypothetical prote
22	39	55.7	1750	2 H64403	ribonucleoside-tri
23	38	54.3	60	2 A48396	ribosomal protein
24	38	54.3	62	2 AH1301	ribosomal protein
25	38	54.3	62	2 AH1673	ribosomal protein
26	38	54.3	205	2 E90094	26S proteasome SU
27	38	54.3	274	2 B60950	apolipoprotein B-1
28	37	52.9	196	2 T00702	hypothetical prote
29	37	52.9	225	2 S41031	hypothetical prote

30 37 52.9 236 2 A84586  
31 37 52.9 289 2 G86403  
32 37 52.9 306 1 A39654  
33 37 52.9 413 2 T32831  
34 37 52.9 500 2 G71633  
35 37 52.9 502 2 T01179  
36 37 52.9 513 2 B86156  
37 37 52.9 749 2 A45687  
38 37 52.9 842 2 T04555  
39 37 52.9 2700 2 D88450  
40 36.5 52.1 235 2 E91097  
41 36.5 52.1 235 2 A85943  
42 36 51.4 116 1 Q0EBHT  
43 36 51.4 116 2 S09523  
44 36 51.4 236 2 AH0157  
45 36 51.4 283 2 G98020

#### ALIGNMENTS

##### RESULT 1

###### THUL

lactotransferrin precursor [validated] - human

N/Alternate names: lactoferrin

C/Species: Homo sapiens (man)

C/Date: 31-Mar-1992 #sequence, revision 21-Nov-1997 #text change 08-Dec-2000

C/Accession: G01394; S11228; A45401; S10324; S15853; S20841; S07160; A61169; A31000; S

R/Cho. V.

submitted to the EMBL Data Library, March 1994

A/Reference number: G06820

A/Accession: G01394

A/Status: preliminary; translated from GB/EMBL/DDBJ

A/Molecule type: mRNA

A/Residues: 1-711 <CHO>

A/Cross-references: EMBL:U07643; NID:G467236; PIDN:AAB60324.1; PID:G467237

R/Rev. M.W. Moloshuk, S.L.; deBoer, H.A.; Pieper, F.R.

Nucleic Acids Res. 18, 5288, 1990

A/Title: Complete nucleotide sequence of human mammary gland lactoferrin.

A/Reference number: S11228; MUID:90384839; PMID:2402455

A/Accession: S11228

A/Molecule type: mRNA

A/Residues: 1-148, 'T', 150-422, 'C', 424-711 <REV>

A/Cross-references: EMBL:X53961; NID:G34415; PIDN:CAA37914.1; PID:G34416

R/Teng, C.T.; Liu, Y.; Yang, N.; Walmer, D.; Panella, T.

Mol. Endocrinol. 6, 1963-1961, 1992

A/Title: Differential molecular mechanism of the estrogen action that regulates lactof

A/Reference number: A45401; MUID:93125571; PMID:1480183

A/Accession: A45401

A/Molecule type: DNA

A/Residues: 1-15 <TEN>

A/Cross-references: GB:S52659; NID:G263311; PIDN:AAB24877.1; PID:G263312

A/Experimental source: placenta

A/Note: sequence extracted from NCBI backbone (NCBIP:122202)

R/Powell, M.J.; Ogden, J.E.

Nucleic Acids Res. 18, 4013, 1990

A/Title: Nucleotide sequence of human lactoferrin cDNA.

A/Reference number: S10324; MUID:90326549; PMID:2374734

A/Accession: S10324

A/Molecule type: mRNA

A/Residues: 3-711 <POW>

A/Cross-references: EMBL:X52941; NID:G34411; PIDN:CAA37116.1; PID:G34412

R/Stowell, K.M.; Rado, T.A.; Funk, W.D.; Tweedie, J.W.

Biochem. J. 276, 349-355, 1991

A/Title: Expression of cloned human lactoferrin in baby-hamster kidney cells.

A/Reference number: S15853; MUID:91264786; PMID:2049066

A/Accession: S15853

A/Status: nucleic acid sequence not shown; not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 20-31 <ST1>

A/Accession: S20841

A/Molecule type: protein

A/Residues: 20-28, 'X', 30-31 <ST2>

R:Rado, T.A.; Wei, X.; Benz Jr., E.J.  
 Blood 70, 989-993, 1987  
 A:Title: Isolation of lactoferrin cDNA from a human myeloid library and expression of mRNA  
 A:Reference number: S07160; MUID:88001031; PMID:3477300  
 A:Accession: S07160  
 A:Molecule type: mRNA  
 A:Cross-references: EMBL:M18642; NID:G186815; PIDN:AAA8665.1; PID:G386855  
 R:Panella, T.J.; Liu, Y.; Huang, A.T.; Teng, C.T.  
 Cancer Res. 51, 3037-3043, 1991  
 A:Title: Polymorphism and altered methylation of the lactoferrin gene in normal leukocytes  
 A:Reference number: A61169; MUID:91235214; PMID:1674448  
 A:Accession: A61169  
 A>Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 3-701, 'SWKPNV' <PAN>  
 A:Experimental source: normal breast tissue  
 R:Metz-Boutigue, M.H.; Jolles, J.; Mazurier, J.; Schoentgen, F.; Legrand, D.; Spik, G.; Eur. J. Biochem. 145, 659-666, 1984  
 A:Title: Human lactotransferrin: amino acid sequence and structural comparisons with other  
 A:Reference number: A31000; MUID:85076667; PMID:6510420  
 A:Accession: A31000  
 A:Molecule type: protein  
 A:Residues: 20-140, 142-169, 171-203, 'L', 205, 'K', 207-208, 'K', 210-385, 'Q', 387-391, 'W', 393-4  
 A:Note: this is the final paper in a series  
 R:Houen, G.; Hoegdall, E.V.; Barkholt, V.; Norskov, L.  
 Eur. J. Biochem. 241, 303-308, 1996  
 A:Title: Lactoferrin: similarity to diamine oxidase and purification by aminohexyl affinity  
 A:Reference number: S74119; MUID:97054624; PMID:8898921  
 A:Accession: S74119  
 A:Molecule type: protein  
 A:Residues: 'G', 23-24, 'R', 26-27, 'XX', 30-32 <HOU>  
 A:Experimental source: neutrophil granulocytes  
 C:Genetics:  
 A:Gene: GDB:LTF  
 A:Cross-references: GDB:119368; OMIM:150210  
 A:Map position: 3c21-3q23  
 C:Superfamily: transferrin; transferrin repeat homology  
 C:Keywords: duplication; glycoprotein; iron binding; milk  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-711/Product: lactotransferrin #status experimental <MAT>  
 F:21-356/Domain: transferrin repeat homology <TRH1>  
 F:360-699/Domain: transferrin repeat homology <TRH2>  
 F:29-65, 39-56, 135-218, 177-193, 190-201, 251-365, 503-697, 595-609/Disulfide bonds: #status e  
 F:157, 438/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F:368-400, 378-391, 425-706, 447-669, 479-554, 513-527, 524-537, 647-652/Disulfide bonds: #stat

A:Residues: 1-511 <PAR>  
 A:Cross-references: GB:AL513382; PIDN:CAD06049.1; PID:G16504016; GSPDB:GN00176  
 C:Genetics:  
 A:Gene: STY3070  
 Query Match 65.7%; Score 46; DB 2; Length 511;  
 Best Local Similarity 66.7%; Pred. No. 2.5;  
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 CFAWKNRKVR 12  
 |||||  
 Db 350 CFAWDMNKAKVR 361  
 |||||  
 RESULT 3  
 Lactoferrin - goat  
 C:Species: Capra aegagrus hircus (domestic goat)  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 07-May-1999  
 C:Accession: JC2323  
 R:Le Provost, F.; Nocard, M.; Guerin, G.; Martin, P.  
 Biochem. Biophys. Res. Commun. 203, 1324-1332, 1994  
 A:Title: Characterization of the goat lactoferrin cDNA: assignment of the relevant lo  
 A:Reference number: JC2323; MUID:94380047; PMID:8093048  
 A:Accession: JC2323  
 A:Molecule type: mRNA  
 A:Residues: 1-708 <LEP>  
 C:Superfamily: transferrin; transferrin repeat homology  
 C:Keywords: duplication; glycoprotein  
 F:359-696/Domain: transferrin repeat homology <TRH2>  
 F:252, 300, 387, 495, 564/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 Query Match 62.9%; Score 44; DB 2; Length 708;  
 Best Local Similarity 54.5%; Pred. No. 7.6;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CFAWKNRKVR 11  
 |||||  
 Db 38 CYWQRRKRL 48  
 |||||  
 RESULT 4  
 E64639  
 3-deoxy-manno-octulosonate cytidyltransferase (EC 2.7.7.38) [similarity] - Helicoba  
 C:Species: Helicobacter pylori  
 C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 28-Jul-2000  
 C:Accession: E64639  
 R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.  
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKel  
 son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey,  
 Nature 388, 539-547, 1997  
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,  
 A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
 A:Reference number: A64520; MUID:97394467; PMID:9252185  
 A:Accession: E64639  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-393 <TOM>  
 A:Cross-references: GB:AE000604; GB:AE000511; NID:G2314086; PIDN:AAD08000.1; PID:G231.  
 C:Genetics:  
 A:Start codon: TTG  
 C:Superfamily: Chlamydia trachomatis 3-deoxy-manno-octulosonate cytidyltransferase  
 C:Keywords: nucleotidyltransferase  
 Query Match 61.4%; Score 43; DB 2; Length 393;  
 Best Local Similarity 45.5%; Pred. No. 6.6;  
 Matches 5; Conservative 6; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 FAWKRNKRVR 12  
 |||||  
 Db 110 FAWKRNKRVR 120  
 |||||

```

RESULT 5
S52107
lactoferrin - sheep (fragment)
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 14-Jul-1995 #sequence_revision 05-Dec-1997 #text_change 05-Dec-1997
C:Accession: S52107
R:Qian, Z.Y.; Jolles, P.; Migliore-Samour, D.; Fiat, A.M.
Biochim. Biophys. Acta 1243: 25-32, 1995
A:Title: Isolation and characterization of sheep lactoferrin, an inhibitor of platelet a
A:Reference number: S52107; MUID:95127729; PMID:7827104
A:Accession: S52107
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-33 <OIA>
C:Superfamily: transferrin; transferrin repeat homology
C:Keywords: duplication

Query Match 58.6%; Score 41; DB 2; Length 33;
Best Local Similarity 45.5%; Pred. No. 1.5;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFAWKNNRKVR 11
|:|:|:|:|:|
Db 19 CYQWQKQKMKL 29

RESULT 6
C84325
hypothetical protein vngi732c [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: C84325
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.;
Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, F.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: C84325
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-584 <STO>
A:Cross-references: GB:AF004437; NID:gl0581192; PIDN:AAG19967.1; GSPDB:GN00138
C:Genetics:
A:Gene: VNGI732C

Query Match 58.6%; Score 41; DB 2; Length 584;
Best Local Similarity 41.7%; Pred. No. 21;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFAWKNNRKVR 12
|:|:|:|:|:|
Db 445 CFTWRKDMERKR 456

RESULT 7
S77802
hypothetical protein MC003 - Mycoplasma capricolum (fragment)
C:Species: Mycoplasma capricolum
C:Date: 09-Oct-1997 #sequence_revision 31-Oct-1997 #text_change 07-Dec-1999
C:Accession: S77802
R:Bork, P.; Ouzounis, C.; Casari, G.; Schneider, R.; Sander, C.; Dolan, M.; Gilbert, W.;
Mol. Microbiol. 16, 955-967, 1995
A:Title: Exploring the Mycoplasma capricolum genome: a minimal cell reveals its physiolo
A:Reference number: S77739; MUID:96059641; PMID:7476192
A:Accession: S77802
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-267 <BOR>
A:Cross-references: ENBL:Z33006
A:Experimental source: AFCC 27343
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994

```

---

```

C:Genetics:
A:Genetic code: SGC3

Query Match 57.1%; Score 40; DB 2; Length 267;
Best Local Similarity 58.3%; Pred. No. 15;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFAWKNNRKVR 12
|:|:|:|:|:|
Db 159 CFGKKNNRQMR 169

RESULT 8
B82921
serine/threonine kinase UU216 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: B82921
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a n
A:Reference number: A82870
A:Accession: B82921
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-386 <GLA>
A:Cross-references: GB:AE002120; GB:AF222894; NID:g6899167; PIDN:AAF30624.1; GSPDB:GNO
A:Experimental source: serovar 3; biovar 1
C:Genetics:
A:Gene: pkn; UU216
A:Genetic code: SGC3

Query Match 57.1%; Score 40; DB 2; Length 386;
Best Local Similarity 54.5%; Pred. No. 22;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFAWKNNRKVR 11
|:|:|:|:|:|
Db 290 CFAWKKNEDNKL 300

RESULT 9
D33876
carcinoembryonic antigen homolog 4 - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 27-Feb-1990 #sequence_revision 27-Feb-1990 #text_change 23-Jul-1999
C:Accession: D33876
R:Kodellja, V.; Lucas, K.; Barnert, S.; von Kleist, S.; Thompson, J.A.; Zimmermann, W.
J. Biol. Chem. 264, 6906-6912, 1989
A:Title: Identification of a carcinoembryonic antigen gene family in the rat. Analysis
A:Reference number: A33876; MUID:89214106; PMID:2708349
A:Accession: D33876
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-114 <KOD>
A:Cross-references: GB:M60026; NID:g203407; PIDN:AAA40911.1; PID:g554427; GB:J04626; G
C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-term

Query Match 55.7%; Score 39; DB 2; Length 114;
Best Local Similarity 60.0%; Pred. No. 10;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 FAWYRGLRKR 11
|:|:|:|:|:|
Db 42 FAWYRGLRKR 51

RESULT 10
AH3147
hypothetical protein Atu4804 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AH3147

```

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell  
; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001  
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; PMID:11743193

A;Accession: AH3147

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-121 <KUR>

A;Cross-references: GB:AE008689; PIDN:RAL45598.1; PID:gl7743317; GSPDB:GN00187

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: Atu4804

A;Map position: linear chromosome

Query Match 55.7%; Score 39; DB 2; Length 121;  
Best Local Similarity 54.3%; Pred. No. 11;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFAWKNMKVR 11

DB 14 CLAWQRRRV 24

RESULT 11

H97451

Pyridoxamine 5'-phosphate oxidase (AF179611) [imported] - Agrobacterium tumefaciens (str

C;Species: Agrobacterium tumefaciens

C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002

C;Accession: H97451

R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,

A.; Liu, P.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A;Reference number: A97359; PMID:11743194

A;Accession: H97451

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-206 <KUR>

A;Cross-references: GB:AE007869; PIDN:AAK86569.1; PID:gl5155733; GSPDB:GN00169

C;Genetics:

A;Gene: AGR\_C\_1381

A;Map position: circular chromosome

C;Superfamily: pyridoxamine-phosphate oxidase

Query Match 55.7%; Score 39; DB 2; Length 206;  
Best Local Similarity 58.3%; Pred. No. 18;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFAWKNMKVR 12

DB 88 CFHWKSLRQVR 99

RESULT 12

AB2670

Pyridoxamine 5'-phosphate oxidase [imported] - Agrobacterium tumefaciens (strain C58, Du

C;Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 01-Feb-2002

C;Accession: AB2670

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; PMID:11743193

A;Accession: AB2670

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-206 <KUR>

A;Cross-references: GB:AE008688; PIDN:AAL41776.1; PID:gl7739129; GSPDB:GN00186

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: pdxH

A;Map position: circular chromosome

C;Superfamily: pyridoxamine-phosphate oxidase

Query Match 55.7%; Score 39; DB 2; Length 206;

Best Local Similarity 58.3%; Pred. No. 18;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFAWKNMKVR 12

DB 88 CFHWKSLRQVR 99

RESULT 13

AG3441

Probable pyridoxamine-phosphate oxidase (BC 1.4.3.5) [imported] - Brucella melitensis

C;Species: Brucella melitensis

C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 03-May-2002

C;Accession: AG3441

R;Delvecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujier, C.; Los, T.; Ivanov

; Mazur, M.; Goltzman, E.; Selkov, E.; Eizer, P.H.; Hagius, S.; O'Callaghan, D.; Let

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A;Title: The genome sequence of the facultative intracellular pathogen Brucella melit

A;Reference number: AB3252; PMID:11756688

A;Accession: AG3441

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-208 <KUR>

A;Cross-references: GB:AE008917; PIDN:AAL52698.1; PID:gl7993525; GSPDB:GN00190

A;Experimental source: strain 16M

C;Genetics:

A;Gene: BME11517

A;Map position: I

C;Superfamily: Pyridoxamine-phosphate oxidase

C;Keywords: oxidoreductase

Query Match 55.7%; Score 39; DB 2; Length 208;

Best Local Similarity 58.3%; Pred. No. 18;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFAWKNMKVR 12

DB 90 CFHWKSLRQVR 101

RESULT 14

T22597

hypothetical protein F53H4.4 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000

C;Accession: T22597

R;Dobson, R.

submitted to the EMBL Data Library, October 1996

A;Reference number: Z19587

A;Accession: T22597

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-275 <WIL>

A;Cross-references: EMBL:Z81089; PIDN:CAB03137.1; GSPDB:GN00028; CESP:F53H4.4

A;Experimental source: clone F53H4

C;Genetics:

A;Gene: CESP:F53H4.4

A;Map position: X

A;Introns: 67/1; 153/1

C;Superfamily: Caenorhabditis elegans hypothetical protein F53H4.4

Query Match

Best Local Similarity 55.7%; Score 39; DB 2; Length 275;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 FAWKENMRKV 12  
|||:||||  
Db 262 FOWKSMRKT 272

RESULT 15  
F90580  
Hypothetical protein MYPV 5500 [imported] - Mycoplasma pulmonis (strain UAB CTIP?)  
C:Species: Mycoplasma pulmonis  
C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001  
C:Accession: F90580  
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;  
Nucleic Acids Res. 29, 2145-2153, 2001  
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm  
A:Reference number: A99512; MUID:21267165; PMID:11353084  
A:Accession: F90580  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-282 <KUR>  
A:Cross-references: GB:AL445566; PID:g14089965; PIDN:CAC13723.1; GSPDB:GN00153  
A:Experimental source: strain UAB CTIP  
C:Genetics:  
A:Gene: MYPV 5500  
A:Genetic code: SGC3

Query Match 55.7%; Score 39; DB 2; Length 282;  
Best Local Similarity 50.0%; Pred. No. 24;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 FAWKENMRKV 11  
|||:||||  
Db 20 FAWQNNIKKI 29

Search completed: February 21, 2003, 06:02:48  
Job time : 10.6047 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:28:06 ; Search time 5.2093 Seconds  
(without alignments)  
95.544 Million cell updates/sec

Title: US-09-743-107b-94

Perfect score: 70

Sequence: 1 CFAMKRNMRKV 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	87.1	711	1 TRFL_HUMAN	P02788 homo sapien
2	44	62.9	708	1 TRFL_CAVDR	Q9tun0 camelus dro
3	44	62.9	708	1 TRFL_CAPHI	Q29477 capra hircu
4	40	57.1	62	1 RL28_THETN	Q8r9u1 thermoanaer
5	40	57.1	267	1 Y125_MYCCA	P53661 mycoplasma
6	40	57.1	695	1 TRFL_HORSE	Q77811 equus cabal
7	39	55.7	675	1 ZG20_XENLA	P18714 xenopus lae
8	39	55.7	707	1 TRFL_MOUSE	P08071 mus musculu
9	39	55.7	1750	1 Y832_METJA	Q58242 methanococ
10	38	54.3	60	1 RL28_EACST	P23374 bacillus st
11	38	54.3	62	1 RL28_LISMO	Q928j2 listeria mo
12	37	52.9	306	1 BUB2_YEAST	P26448 saccharomyc
13	37	52.9	500	1 TLCE_RICPR	C05962 rickettsia
14	37	52.9	749	1 VP4_ROTGA	Q04916 rotavirus (
15	36	51.4	116	1 MERT_SALTI	P04336 salmonella
16	36	51.4	365	1 LA34_HUMAN	P30453 homo sapien
17	36	51.4	398	1 YK13_CABEL	P34337 caenorhabdi
18	36	51.4	428	1 SYH_CHLMU	Q9pj99 chlamydia m
19	35	50.0	421	1 YDT4_SCHPO	O14209 schizosacch
20	35	50.0	480	1 YOS1_CABEL	C09309 caenorhabdi
21	35	50.0	536	1 YEN1_SCHPO	O13695 schizosacch
22	35	50.0	783	1 YNR2_CABEL	Q21988 caenorhabdi
23	34.5	49.3	250	1 PCEA_MOUSE	P20489 mus musculu
24	34.5	49.3	256	1 YQEC_ECOLI	Q46809 escherichia
25	34	48.6	67	1 ATPE_SCHPO	P87316 schizosacch
26	34	48.6	275	1 IL2A_BOVIN	P12342 bos taurus
27	34	48.6	275	1 IL2A_SHEEP	P26898 ovis aries
28	34	48.6	369	1 SP11_MYXVL	P12393 myxoma viru
29	34	48.6	420	1 YO20_BACHD	Q9ka70 bacillus ha
30	34	48.6	428	1 SVH_CHLTR	O84547 chlamydia t
31	34	48.6	485	1 GLGA_BACST	C08328 bacillus st
32	34	48.6	485	1 MAOX_HUMAN	P48163 homo sapien
33	34	48.6	572	1 MAOX_RAT	P13697 rattus norv

#### ALIGNMENTS

##### RESULT 1

```

TRFL_HUMAN
ID TRFL_HUMAN STANDARD; PRT; 711 AA.
AC P02788; Q16780; Q16786; Q16789; Q00756; Q9H1Z3; Q96K24;
AC Q96K25;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lactoferrin precursor (Lactoferrin) [Contains: Lactoferrin A;
DE Lactoferrin B; Lactoferrin C].
GN LTF OR Lf.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Eutelestomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC TISSUE=Mammary gland;
RX MEDLINE=9038483; PubMed=2402455;
RA Rey M.W., Woloshuk S.L., de Boer H.A., Pieper F.R.;
RT "Complete nucleotide sequence of human mammary gland lactoferrin.";
RL Nucleic Acids Res. 18:5288-5288(1990).
RN [2]
RC TISSUE=Mammary gland;
RP SEQUENCE FROM N.A.
RA Cho Y.Y.;
RT Thesis (1994), Genetic Engineering Research Institute / Taejon, Korea.
RN [3]
RC TISSUE=Mammary gland;
RP SEQUENCE FROM N.A.
RA Conneely O.M.;
RT Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.
RN [4]
RC TISSUE=Mammary gland;
RP SEQUENCE FROM N.A.
RA Liang Q., Jimenez-Flores R., Richardson T.;
RT "Molecular cloning and sequence analysis of human lactoferrin.";
RL Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.
RN [5]
RC TISSUE=Mammary gland;
RP SEQUENCE FROM N.A.
RA Wei X., Han J., Rado T.A.;
RT "Human neutrophil lactoferrin coding and 5' flanking region DNA
sequences.";
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
RN [6]
RC TISSUE=Mammary gland;
RP SEQUENCE FROM N.A.
RA Cheng H., Chen X., Huan L.;
RT "cDNA cloning and sequence analysis of human lactoferrin.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [7]
RC TISSUE=Prostate;
RP SEQUENCE FROM N.A.
RA Strauberg R.;
RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [8]

```

```

34 34 48.6 1025 1 P021_NASVI
35 34 48.6 1064 1 CYA4_RAT
36 34 48.6 1179 1 ATX1_ARATH
37 34 48.6 1293 1 XPC_DROME
38 34 48.6 1574 1 RPOC_AQUAE
39 34 48.6 1576 1 RPOC_AQUAE
40 34 48.6 1592 1 YHD5_YEAST
41 34 48.6 1696 1 ITN2_HUMAN
42 34 48.6 2273 1 ABCR_HUMAN
43 34 48.6 4568 1 DYHB_CHLRE
44 33.5 47.9 323 1 ATB2_DROME
45 33 47.1 49 1 YIM9_BPH1

```

```

Q03278 nasonia vit
P26770 rattus norv
Q91502 arabidopsis
Q24595 drosophila
O67763 aquifex aeo
Q9x6Y2 aquifex pyr
P38735 saccharomyc
Q9nm33 homo sapien
P78363 homo sapien
Q39565 chlamydomon
Q24048 drosophila
P10432 bacterioph

```

RP SEQUENCE OF 3-711 FROM N.A.  
RC TISSUE=Mammary gland;  
RX MEDLINE=90326549; PubMed=2374734;  
RA Powell M.J., Ogden J.E.;  
RT "Nucleotide sequence of human lactoferrin cDNA.";  
RL Nucleic Acids Res. 18:4013-4013(1990).  
RN [9]  
RP SEQUENCE OF 20-711.  
RX MEDLINE=85076667; PubMed=6510420;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Schoentgen F.,  
R LeGrand D., Spik G., Montreuil J., Jolles P.;  
RT "Human lactotransferrin: amino acid sequence and structural  
RT comparisons with other transferrins.";  
RL Eur. J. Biochem. 145:659-666(1984).  
RN [10]  
RP PRELIMINARY SEQUENCE OF 20-73; 134-171; 257-278 AND 347-530.  
RX MEDLINE=82046817; PubMed=6794640;  
RA Metz-Boutigue M.-H., Mazurier J., Jolles J., Spik G., Montreuil J.,  
R Jolles P.;  
RT "The present state of the human lactotransferrin sequence. Study and  
RT alignment of the cyanogen bromide fragments and characterization of  
RT N- and C-terminal domains.";  
RL Biochim. Biophys. Acta 670:243-254(1981).  
RN [11]  
RP SEQUENCE OF 609-711.  
RX MEDLINE=82262043; PubMed=7049727;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,  
R Jolles P.;  
RT "An 88 amino acid long C-terminal sequence of human  
RT lactotransferrin.";  
RL FEBS Lett. 142:107-110(1982).  
RN [12]  
RP SEQUENCE OF 436-711 FROM N.A.  
RX MEDLINE=8801031; PubMed=3477300;  
RA Rado T.A., Wei X., Benz E.J. Jr.;  
RT "Isolation of lactoferrin cDNA from a human myeloid library and  
RT expression of mRNA during normal and leukemic myelopoiesis.";  
RL Blood 70:989-993(1987).  
RN [13]  
RP SEQUENCE OF 237-711 FROM N.A.  
RA McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,  
R Nian M., Parnell L., Pedria N., Ansari A., Mardis E., Schutz K.,  
R Guoj L., la Bastide M., Kaplan N., Greco T., Touchman J., Muzny D.,  
R Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M., Porcel B.M.,  
R Dragan Y., Giacalone J., Pae A., Powell E., Solinsky K.A., Desilva U.,  
R Diaz-Perez S., Zhou X., Yu Y., Watanabe M., Doggett N., Garcia D.,  
R Segripanti J.L.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
RN [14]  
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS), AND REVISIONS.  
RX MEDLINE=90064528; PubMed=2585506;  
RA Anderson B.F., Baker H.M., Norris G.E., Rice D.W., Baker E.N.;  
RT "Structure of human lactoferrin: crystallographic structure analysis  
RT and refinement at 2.8-A resolution.";  
RL J. Mol. Biol. 209:711-734(1989).  
RN [15]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RA Haridas M., Anderson B.F., Baker E.N.;  
RT "Structure of human diferric lactoferrin refined at 2.2-A  
RT resolution.";  
RL Acta Crystallogr. D 51:629-646(1995).  
RN [16]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 20-353.  
RX MEDLINE=97156796; PubMed=9003186;  
RA Nicholson H., Anderson B.F., Bland T., Shewry S.C., Tweedie J.W.,  
R Baker E.N.;  
RT "Mutagenesis of the histidine ligand in human lactoferrin: iron  
RT binding properties and crystal structure of the histidine-  
RT 253--methionine mutant.";  
RL Biochemistry 36:341-346(1997).  
RN [17]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX MEDLINE=99190892; PubMed=10089347;  
RN Sun X.L., Baker H.M., Shewry S.C., Jameson G.B., Baker E.N.;  
RT "Structure of recombinant human lactoferrin expressed in Aspergillus  
RT awamori.";  
RL Acta Crystallogr. D 55:403-407(1999).  
RN [18]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=99192677; PubMed=1009508;  
RA Jameson G.B., Anderson B.F., Norris G.E., Thomas D.H., Baker E.N.;  
RT "Structure of human apolactoferrin at 2.0-A resolution. Refinement  
RT and analysis of ligand-induced conformational change.";  
RL Acta Crystallogr. D 54:1319-1335(1998).  
RN [19]  
RP CHARACTERIZATION OF LACTOFERROXINS.  
RX MEDLINE=91166329; PubMed=1369293;  
RA Tani F., Iio K., Chiba H., Yoshikawa M.;  
RT "Isolation and characterization of opioid antagonist peptides derived  
RT from human lactoferrin.";  
RL Agric. Biol. Chem. 54:1803-1810(1990).  
RN [20]  
RP VARIANTS THR-30 AND ARG-48.  
RX PubMed=9873069;  
RA Klintworth G.K., Sommer J.R., Obrian G., Han L., Ahmed M.N.,  
R Qumsiyeh M.B., Lin P.-Y., Basti S., Reddy M.K., Kanai A., Hotta Y.,  
R Sugar J., Kumaramanickavel G., Munier F., Schorderet D.F.,  
R El Matiri L., Iwata F., Kaiser-Kupfer M., Nagata M., Nakayasu K.,  
R Hejtmancik J.F., Teng C.T.;  
RT "Familial subepithelial corneal amyloidosis (gelatinous drop-like  
RT corneal dystrophy): exclusion of linkage to lactoferrin gene.";  
RL Mol. Vision 4:31-32(1998).  
CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
CC OF AN ANION, USUALLY BICARBONATE.  
CC -1- FUNCTION: LACTOFERROXINS A, B AND C HAVE OPIOID ANTAGONIST  
CC ACTIVITY. LACTOFERRIN A SHOWS PREFERENCE FOR MU-RECEPTORS, WHILE  
CC LACTOFERRIN B AND C HAVE SOMEWHAT HIGHER DEGREES OF PREFERENCE  
CC FOR KAPPA-RECEPTORS THAN FOR MU-RECEPTORS.  
CC -1- SUBUNIT: MONOMER.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR ENBL; X53961; CAA37914.1; -  
DR ENBL; U07643; AAB60324.1; -  
DR ENBL; M93150; AAA36159.1; -  
DR ENBL; M83202; AAA59511.1; -  
DR ENBL; M83205; AAA58656.1; -  
DR ENBL; M18642; AAA86665.1; -  
DR ENBL; AF332168; AAG48753.1; -  
DR ENBL; BC015822; AAH15822.1; -  
DR ENBL; BC015823; AAH15823.1; -  
DR ENBL; M73700; AAA59479.1; -  
DR ENBL; X52941; CAA37116.1; -  
DR ENBL; U95626; AAB57795.1; -  
DR PIR; S11228; TFHUL.  
DR PDB; 1LCF; 31-AUG-94.  
DR PDB; 1LCT; 31-OCT-93.  
DR PDB; 1LFG; 31-JUL-94.  
DR PDB; 1LFF; 31-OCT-93.  
DR PDB; 1LFI; 31-OCT-93.  
DR PDB; 1LGB; 31-AUG-94.  
DR PDB; 1LGC; 31-AUG-94.  
DR PDB; 1BKA; 08-NOV-96.  
DR PDB; 1DSN; 08-MAR-96.  
DR PDB; 1HSE; 12-MAR-97.  
DR PDB; 1VFD; 21-APR-97.

Query Match 87.1%; Score 61; DB 1; Length 711;  
 Best Local Similarity 83.3%; Pred.No. 0.0015;  
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFAWKNMKV 12  
 DB 39 CFAWQRMKVR 50

## RESULT 2

TRFL\_CAMDR STANDARD; PRT; 708 AA.  
 AC OPTUNO; QWZSS;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin).  
 GN LTF.  
 OS Camelus dromedarius (Dromedary) (Arabian camel).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.  
 OX NCBI\_TaxID=9838;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Somali; TISSUE=Lactating mammary gland;  
 RA Kappeler S.R.; Ackermann M.; Farah Z.; Puhon Z.;  
 RT "Sequence analysis of camel (Camelus dromedarius) lactoferrin.";  
 RL Int. Dairy J. 9:481-486 (1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Paramasivam M.; Srinivasan A.; Singh R.; Sahani M.S.; Singh T.P.;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 OF AN ANION, USUALLY BICARBONATE (BY SIMILARITY).  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AJ131674; CAB53387.1; -;  
 DR EMBL; AF165879; AAF82241.1; -;  
 DR HSP; 077811; 1BLX.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin.2.  
 DR PRINTS; PRO0422; TRANSFERRIN.  
 DR SMART; SM00094; TR\_FER; 2.  
 DR PROSITE; PS00205; TRANSFERRIN\_1; 2.  
 DR PROSITE; PS00206; TRANSFERRIN\_2; 2.  
 DR PROSITE; PS00207; TRANSFERRIN\_3; 2.  
 KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 KW Signal.  
 FT SIGNAL. 1 19 BY SIMILARITY.  
 FT CHAIN 20 708 LACTOTRANSFERRIN.  
 FT REPEAT 20 363 1.  
 FT REPEAT 364 708 2.  
 FT DISULFID 28 64 BY SIMILARITY.  
 FT DISULFID 38 55 BY SIMILARITY.  
 FT DISULFID 134 217 BY SIMILARITY.  
 FT DISULFID 176 192 BY SIMILARITY.  
 FT DISULFID 189 200 BY SIMILARITY.  
 FT DISULFID 250 264 BY SIMILARITY.  
 FT DISULFID 367 399 BY SIMILARITY.

FT DISULFID 377 390 BY SIMILARITY.  
 FT DISULFID 424 703 BY SIMILARITY.  
 FT DISULFID 444 666 BY SIMILARITY.  
 FT DISULFID 476 551 BY SIMILARITY.  
 FT DISULFID 500 694 BY SIMILARITY.  
 FT DISULFID 510 524 BY SIMILARITY.  
 FT DISULFID 521 534 BY SIMILARITY.  
 FT DISULFID 592 606 BY SIMILARITY.  
 FT DISULFID 644 649 BY SIMILARITY.  
 FT METAL 79 79  
 FT METAL 111 111 IRON 1 (BY SIMILARITY).  
 FT METAL 211 211 IRON 1 (BY SIMILARITY).  
 FT METAL 272 272 IRON 1 (BY SIMILARITY).  
 FT METAL 414 414 IRON 2 (BY SIMILARITY).  
 FT METAL 452 452 IRON 2 (BY SIMILARITY).  
 FT METAL 545 545 IRON 2 (BY SIMILARITY).  
 FT METAL 614 614 IRON 2 (BY SIMILARITY).  
 FT BINDING 140 140 ANION (BY SIMILARITY).  
 FT BINDING 482 482 ANION (BY SIMILARITY).  
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .)  
 FT CONFLICT 261 261 F -> S (IN REF. 2).  
 FT CONFLICT 304 304 G -> A (IN REF. 2).  
 FT CONFLICT 330 330 S -> P (IN REF. 2).  
 FT CONFLICT 492 494 LLS -> PLF (IN REF. 2).  
 FT CONFLICT 506 506 L -> F (IN REF. 2).  
 FT CONFLICT 609 609 A -> P (IN REF. 2).  
 FT CONFLICT 642 642 R -> Q (IN REF. 2).  
 SQ SEQUENCE 708 AA; 77211 MW; 0B0C175A0B69D430 CRC64;

Query Match 62.9%; Score 44; DB 1; Length 708;

Best Local Similarity 58.3%; Pred.No. 1.9;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFAWKNMKV 12  
 DB 38 CAQWQRMKVR 49

## RESULT 3

TRFL\_CAMPH STANDARD; PRT; 708 AA.  
 ID TRFL\_CAMPH  
 AC Q29477; Q29479;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin).  
 GN LTF.  
 OS Capra hircus (Goat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 CC Bovidae; Caprinae; Capra.  
 OX NCBI\_TaxID=9925;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Lee T.; Yu S.; Kim S.; Lee K.; Yu D.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RX MEDLINE=9438004; Pubmed=8093048;  
 RA le Provost F.; Nocard M.; Guerin G.; Martin P.;  
 RT "Characterization of the goat lactoferrin cDNA. Assignment of the  
 relevant locus to bovine U12 syntenic group."  
 RL Biochem. Biophys. Res. Commun. 203:1324-1332(1994).  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 OF AN ANION, USUALLY BICARBONATE.  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Secreted.

CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U53857; AAA57958.1; --  
 CC EMBL; X78902; CAA55517.1; --  
 CC HSP; O77698; ICE2.  
 CC InterPro; IPR001156; Transferrin.  
 CC Pfam; PF00405; transferrin; 2.  
 CC PRINTS; PR00422; TRANSFERRIN.  
 CC SMART; SMC0094; TR\_FER; 2.  
 CC PROSITE; PS00205; TRANSFERRIN\_1; 2.  
 CC PROSITE; PS00206; TRANSFERRIN\_2; 2.  
 CC DR PROSITE; PS00207; TRANSFERRIN\_3; 2.  
 CC KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 CC Signal.  
 CC FT SIGNAL 1 19  
 CC FT CHAIN 20 708  
 CC FT REPEAT 20 363  
 CC FT REPEAT 364 708  
 CC FT DISULFID 28 64 BY SIMILARITY.  
 CC FT DISULFID 38 55 BY SIMILARITY.  
 CC FT DISULFID 134 217 BY SIMILARITY.  
 CC FT DISULFID 176 192 BY SIMILARITY.  
 CC FT DISULFID 189 200 BY SIMILARITY.  
 CC FT DISULFID 250 264 BY SIMILARITY.  
 CC FT DISULFID 357 399 BY SIMILARITY.  
 CC FT DISULFID 377 390 BY SIMILARITY.  
 CC FT DISULFID 424 703 BY SIMILARITY.  
 CC FT DISULFID 444 666 BY SIMILARITY.  
 CC FT DISULFID 476 551 BY SIMILARITY.  
 CC FT DISULFID 500 694 BY SIMILARITY.  
 CC FT DISULFID 510 524 BY SIMILARITY.  
 CC FT DISULFID 521 534 BY SIMILARITY.  
 CC FT DISULFID 592 606 BY SIMILARITY.  
 CC FT DISULFID 644 649 BY SIMILARITY.  
 CC FT METAL 79 79 IRON 1 (BY SIMILARITY).  
 CC FT METAL 111 111 IRON 1 (BY SIMILARITY).  
 CC FT METAL 211 211 IRON 1 (BY SIMILARITY).  
 CC FT METAL 272 272 IRON 1 (BY SIMILARITY).  
 CC FT METAL 414 414 IRON 2 (BY SIMILARITY).  
 CC FT METAL 452 452 IRON 2 (BY SIMILARITY).  
 CC FT METAL 545 545 IRON 2 (BY SIMILARITY).  
 CC FT METAL 614 614 IRON 2 (BY SIMILARITY).  
 CC FT BINDING 140 140 ANION (BY SIMILARITY).  
 CC FT BINDING 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 495 495 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 564 564 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CONFLICT 56 56 I -> V (IN REF. 2).  
 CC FT CONFLICT 88 88 L -> R (IN REF. 2).  
 CC FT CONFLICT 124 124 Q -> K (IN REF. 2).  
 CC FT CONFLICT 154 154 F -> P (IN REF. 2).  
 CC FT CONFLICT 304 304 S -> R (IN REF. 2).  
 CC FT CONFLICT 414 414 D -> G (IN REF. 2).  
 CC SQ SEQUENCE 708 AA; 77358 MW; F2EDA3C83539960D CRC64;  
 CC  
 CC Query Match 62.9%; Score 44; DB 1; Length 708;  
 CC Best Local Similarity 54.5%; Pred. No. 1.9;  
 CC Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 CC  
 CC QY 1 CFAWKNRKRV 11  
 CC : : : : :  
 CC Db 38 CQWQREMRKL 48

RESULT 4  
 RL28\_THETN STANDARD; PRT; 62 AA.  
 ID RL28\_THETN  
 AC Q8R9U1  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 50S ribosomal protein L28.  
 GN RPB OR TTE1495.  
 OS Thermoanaerobacter tengcongensis.  
 OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;  
 OC Thermoanaerobacteriaceae; Thermoanaerobacter.  
 OX NCBI\_TaxID=119072;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MB4 / JCM 11007;  
 RX MEDLINE=21992816; PubMed=11997336;  
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;  
 RA "A complete sequence of T. tengcongensis genome."  
 RL Genome Res. 12:689-700(2002).  
 CC -1- SIMILARITY: BELONGS TO THE L28P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AE013107; AAM24713.1; --  
 CC KW Ribosomal protein; Complete proteome.  
 CC SQ SEQUENCE 62 AA; 7037 MW; CAADB605C81D495D CRC64;  
 CC  
 CC Query Match 57.1%; Score 40; DB 1; Length 62;  
 CC Best Local Similarity 77.8%; Pred. No. 0.81;  
 CC Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 CC QY 4 WKFNMRKVR 12  
 CC : : : : :  
 CC Db 28 WKFNMRKVR 36  
 CC  
 CC RESULT 5  
 ID Y125\_MYCCA STANDARD; PRT; 267 AA.  
 AC P53661;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Hypothetical protein (fragment).  
 OS Mycoplasma capricolum.  
 OC Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;  
 OC Entomoplasmataceae.  
 OX NCBI\_TaxID=2095;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 27343 / KID;  
 RX MEDLINE=96059641; PubMed=7476192;  
 RA Bork P., Ouzounis C., Casari G., Schneider R., Sander C.,  
 RA Dolan M., Gilbert W., Gillevet P.M.;  
 RA "Exploring the Mycoplasma capricolum genome: a minimal cell reveals  
 RT its physiology."  
 RL Mol. Microbiol. 16:955-967(1995).  
 CC -1- SIMILARITY: BELONGS TO THE COF/YBHA/YIDA/YIGL (E.COLI) / YCSE/YXEH  
 CC (B.SUBTILIS) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

File # : 2009-0076  
Date Rec'd : 08/06/2009

[2]  
 RN RP SEQUENCE OF 85-613 FROM N.A.  
 RX MEDLINE=90040698; PubMed=2509712;  
 RA Nicfield W., El-Baradi T., Mentzel H., Pieler T., Koester M.,  
 RA Poeling A., Knoechel W.;  
 RT J. Mol. Biol. 208:639-659(1989).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X82643; CAA57965.1; --  
 DR PIR; S06565; S06565.  
 DR HSP; P08046; IALI.  
 DR TRANSFAC; T02366; --  
 DR InterPro; IPR000822; Znf C2H2.  
 DR Pfam; PF00096; zf-C2H2; 18.  
 DR PRINTS; PR00048; ZNCFINGER.  
 DR ProDom; PD000003; Znf C2H2; 5.  
 DR SMART; SM00355; Znf C2H2; 18.  
 DR PROSITE; PS00028; ZINC FINGER C2H2 1; 17.  
 DR PROSITE; PS0157; ZINC FINGER C2H2 2; 18.  
 KW Zinc-finger; Metal-binding; DNA-binding; Repeat.  
 FT ZN FING 62 84  
 FT ZN FING 90 112  
 FT ZN FING 118 140  
 FT ZN FING 146 168  
 FT ZN FING 174 196  
 FT ZN FING 202 224  
 FT ZN FING 257 279  
 FT ZN FING 286 308  
 FT ZN FING 344 366  
 FT ZN FING 373 395  
 FT ZN FING 424 446  
 FT ZN FING 452 474  
 FT ZN FING 507 529  
 FT ZN FING 535 557  
 FT ZN FING 563 585  
 FT ZN FING 591 613  
 FT ZN FING 619 642  
 SQ SEQUENCE 675 AA; 77116 MW; 033094852C1FCF39 CRC64;  
 Query Match 55.7%; Score 39; DB 1; Length 675;  
 Best Local Similarity 85.7%; Pred. No. 15;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 FAWKENM 8  
 Db 600 FAWKENL 606  
 RESULT 8  
 ID TRFL MOUSE STANDARD; PRT; 707 AA.  
 AC P08071; P70690; Q61799; Q22222;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin).  
 GN LTF.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Uterus;  
 RX MEDLINE=87280033; PubMed=3611056;

RA Pentecost B.T., Teng C.T.;  
 RT "Lactotransferrin is the major estrogen inducible protein of mouse  
 RT uterine secretions.";  
 RL J. Biol. Chem. 262:10134-10139(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA TISSUE=Uterus;  
 RA Moriishi K.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 1-14 FROM N.A.  
 RX MEDLINE=92042099; PubMed=1939212;  
 RA Liu Y., Teng C.T.;  
 RT "Characterization of estrogen-responsive mouse lactoferrin promoter.";  
 RL J. Biol. Chem. 266:21880-21885(1991).  
 CC -I- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY BICARBONATE.  
 CC -I- SUBUNIT: MONOMER.  
 CC -I- SUBCELLULAR LOCATION: Secreted.  
 CC -I- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -I- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; J03298; AAA40525.1; --  
 DR EMBL; D88510; BAAL3633.1; --  
 DR EMBL; BC006904; AAH06904.1; --  
 DR EMBL; M74778; AAA39427.1; --  
 DR PIR; A28438; A28438.  
 DR HSP; P02789; ICB6.  
 DR MGD; MGI:96837; Ltf.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 2.  
 DR PRINTS; PR00422; TRANSFERRIN.  
 DR SMART; SM00094; TR\_FER; 2.  
 DR PROSITE; PS00205; TRANSFERRIN\_1; 1.  
 DR PROSITE; PS00206; TRANSFERRIN\_2; 2.  
 DR PROSITE; PS00207; TRANSFERRIN\_3; 2.  
 KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 KW Signal.  
 FT SIGNAL 1 19 BY SIMILARITY.  
 FT CHAIN 20 707 LACTOTRANSFERRIN.  
 FT REPEAT 20 357 1.  
 FT REPEAT 358 707 2.  
 FT D-SULFID 27 63 BY SIMILARITY.  
 FT D-SULFID 37 54 BY SIMILARITY.  
 FT D-SULFID 133 216 BY SIMILARITY.  
 FT D-SULFID 175 191 BY SIMILARITY.  
 FT D-SULFID 188 199 BY SIMILARITY.  
 FT D-SULFID 249 263 BY SIMILARITY.  
 FT D-SULFID 366 398 BY SIMILARITY.  
 FT D-SULFID 376 389 BY SIMILARITY.  
 FT D-SULFID 423 702 BY SIMILARITY.  
 FT D-SULFID 443 665 BY SIMILARITY.  
 FT D-SULFID 475 550 BY SIMILARITY.  
 FT D-SULFID 499 693 BY SIMILARITY.  
 FT D-SULFID 509 523 BY SIMILARITY.  
 FT D-SULFID 520 533 BY SIMILARITY.  
 FT D-SULFID 591 605 BY SIMILARITY.  
 FT D-SULFID 643 648 BY SIMILARITY.  
 FT METAL 78 78 BY SIMILARITY.  
 FT METAL 110 110 IRON 1 (BY SIMILARITY).  
 FT METAL 110 110 IRON 1 (BY SIMILARITY).

```

FT METAL 210 210 IRON 1 (BY SIMILARITY)
FT METAL 271 271 IRON 1 (BY SIMILARITY)
FT METAL 413 413 IRON 2 (BY SIMILARITY)
FT METAL 451 451 IRON 2 (BY SIMILARITY)
FT METAL 544 544 IRON 2 (BY SIMILARITY)
FT METAL 613 613 IRON 2 (BY SIMILARITY)
FT BINDING 139 139 ANION (POTENTIAL)
FT BINDING 481 481 ANION (POTENTIAL)
FT CARBOHYD 118 118 N-LINKED (GLCNAC)
FT CARBOHYD 494 494 N-LINKED (GLCNAC)
FT CONFLICT 1 2 MR -> IOG (IN REF. 1)
FT CONFLICT 25 25 R -> Q (IN REF. 2)
FT CONFLICT 82 82 S -> L (IN REF. 2)
FT CONFLICT 359 359 M -> T (IN REF. 2)
FT CONFLICT 382 382 A -> D (IN REF. 1)
FT CONFLICT 449 449 E -> G (IN REF. 2)
FT CONFLICT 629 629 L -> V (IN REF. 1)
SQ SEQUENCE 707 AA; 77865 MW; F26AE0340A4C19A8 CRC64;

Query Match 55.7%; Score 39; DB 1; Length 707;
Best Local Similarity 54.5%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFAWKNNRKY 11
DB 37 CLRWQNNRKY 47

RESULT 9
Y832 METJA STANDARD; PRT; 1750 AA.
AC Q58242;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0832 [Contains: Mja rnr-1 intein; Mja rnr-2 intein].
GN MJ0832.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999, PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kervatage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii."
RL Science 273:1058-1073(1996).
CC -!- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTNIN) FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).
CC -!- SIMILARITY: WEAK IN THE C-TERMINAL, TO M.JANNASCHII MJ0885.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67527; AAB98034.1;
DR HSSP; P17255; 1DFA.
DR TIGR; MJ0832;

```

```

DR InterPro; IPR005144; ATP.
DR InterPro; IPR003586; Hedgehog hintc.
DR InterPro; IPR003587; Hedgehog hintn.
DR InterPro; IPR002203; intein.
DR Pfam; PF03477; ATP-cone; 2.
DR PRINTS; PRO0379; INTEIN.
DR SMART; SM00306; HincC; 2.
DR PROSITE; PS00818; INTEIN_C_TER; 2.
DR PROSITE; PS00819; INTEIN_ENDONUCLEASE; 2.
DR PROSITE; PS00817; INTEIN_N_TER; 2.
KW Hypothetical protein; Autocatalytic cleavage; Protein splicing;
FT CHAIN 1 337 HYPOTHETICAL PROTEIN MJ0832, 1ST PART (POTENTIAL).
FT CHAIN 338 790 MJA RNR-1 INTEIN (POTENTIAL).
FT CHAIN 791 1058 HYPOTHETICAL PROTEIN MJ0832, 2ND PART (POTENTIAL).
FT CHAIN 1059 1591 MJA RNR-2 INTEIN (POTENTIAL).
FT CHAIN 1592 1750 HYPOTHETICAL PROTEIN MJ0832, 3RD PART (POTENTIAL).
SQ SEQUENCE 1750 AA; 203279 MW; 6D6893B770A25359 CRC64;

Query Match 55.7%; Score 39; DB 1; Length 1750;
Best Local Similarity 54.5%; Pred. No. 39;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 FAWKNNRKY 12
DB 718 YAWKNNRKLK 728

RESULT 10
RL28 BACST STANDARD; PRT; 60 AA.
ID RL28 BACST STANDARD; PRT; 60 AA.
AC P23374;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L28.
GN RPMB.
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillales; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE.
RX MEDLINE=92075758; PubMed=1742360;
RA Kruff V., Kapp U., Wittmann-Liebold B.;
RT "Characterization and primary structure of proteins L28, L33 and L34 from Bacillus stearothermophilus ribosomes."
RL Biochimie 73:855-860(1991).
CC -!- SIMILARITY: BELONGS TO THE L28P FAMILY OF RIBOSOMAL PROTEINS.
DR FIR; A48396; A48396.
DR InterPro; IPR001383; Ribosomal_L28.
DR Pfam; PF00830; Ribosomal_L28; 1.
DR TIGRPFAMs; TIGR00009; L28; 1.
FT INIT MET 0
FT SEQUENCE 60 AA; 6810 MW; 2AD9161CD60B92F4 CRC64;

Query Match 54.3%; Score 38; DB 1; Length 60;
Best Local Similarity 56.7%; Pred. No. 1.8;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 WKNNRKY 12
DB 27 WKANLQKVR 35

RESULT 11
RL28 LISMO STANDARD; PRT; 62 AA.
ID RL28 LISMO STANDARD; PRT; 62 AA.

```

AC Q02AJ2;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 50S ribosomal protein L28.  
 GN RPMB OR LM01816 OR LIN1930.  
 OS Listeria monocytogenes, and  
 OS Listeria innocua.  
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 OX NCBI\_TaxID=1639, 1642;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES: L. monocytogenes, and L. innocua;  
 RC STRAIN=BGD-9 / Serovar 1/2a, and CLIP 11262 / Serovar 6a;  
 RX MEDLINE=21537279; PubMed=11679669;  
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,  
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,  
 RA Charbit A., Chetoui F., Couve E., de Daruvar A., Dehoux P.,  
 RA Domant E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
 RA Etienne K.-D., Fsihi H., Garcia-del Portillo P., Garrido P.,  
 RA Gautier L., Goshel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
 RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkak G.,  
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Neelari H.,  
 RA Nardonek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,  
 RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,  
 RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;  
 RT "Comparative genomics of *Listeria* species.";  
 RL Science 294:849-852(2001).  
 CC -1- SIMILARITY: BELONGS TO THE L28P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AL591981; CAC99894.1; -;  
 DR EMBL; AL596170; CAC97160.1; -;  
 DR ListL1st; LIN01930; -;  
 DR ListL1st; LM001816; -;  
 DR InterPro; IPR001383; Ribosomal\_L28.  
 DR Pfam; PF00830; Ribosomal\_L28; 1.  
 DR TIGRFAMs; TIGR00009; L28; 1.  
 KW Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 62 AA; 6991 MW; AA43DE039213C562 CRC64;  
 Query Match 54.3%; Score 38; DB 1; Length 62;  
 Best Local Similarity 66.7%; Pred. No. 1.9;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 WKNNRKYR 12  
 DB 29 WKANLQVR 37  
 RESULT 12  
 BUB2 YEAST  
 ID BUB2 YEAST STANDARD; PRT; 306 AA.  
 AC P26448;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Mitotic checkpoint protein BUB2 (Cell cycle arrest protein BUB2).  
 GN BUB2 OR YNR055C OR YMR796.08C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288c;

RX MEDLINE=91330299; PubMed=1651171;  
 RA Hoyt M.A., Totis L., Roberts B.T.;  
 RT "S. cerevisiae genes required for cell cycle arrest in response to  
 RT loss of microtubule function.";  
 RL Cell 66:507-517(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288c / AB972;  
 RA Devlin K., Churcher C.M., Barrell B.G., Rajandream M.A.;  
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP CHARACTERIZATION.  
 RX MEDLINE=21385309; PubMed=11493673;  
 RA Lee S.E., Jensen S., Frenz L.M., Johnson A.L., Pesquet D.,  
 RA Johnston L.H.;  
 RT "The BUB2-dependent mitotic pathway in yeast acts every cell cycle and  
 RT regulates cytokinesis.";  
 RL J. Cell Sci. 114:2345-2354(2001).  
 CC -1- FUNCTION: Part of a checkpoint which monitors spindle integrity  
 CC and prevents premature exit from mitosis. This cell-cycle arrest  
 CC depends upon inhibition of the G-protein Tem1 by the BFA1/BUB2  
 CC complex.  
 CC -1- SUBUNIT: Interacts with BFA1.  
 CC -1- SUBCELLULAR LOCATION: Spindle poles.  
 CC -1- SIMILARITY: TO S. POMBE CDC16.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M64706; AAA16885.1; -;  
 DR EMBL; Z49703; CAA89765.1; -;  
 DR FIR; A39654; A39654.  
 DR SGD; S0004659; BUB2.  
 DR InterPro; IPR00195; RabGAP\_TBC.  
 DR Pfam; PF00566; TBC; 1.  
 DR SMART; SM00164; TBC; 1.  
 KW Cell cycle; Mitosis.  
 SQ SEQUENCE 305 AA; 35027 MW; A1DDBF548E81EA3 CRC64;  
 Query Match 52.9%; Score 37; DB 1; Length 306;  
 Best Local Similarity 50.0%; Pred. No. 15;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 CFAWKNMRK 10  
 DB 108 CFAWQQR 117  
 RESULT 13  
 TLCE\_RICPR  
 ID TLCE\_RICPR STANDARD; PRT; 500 AA.  
 AC Q059Z;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE ADP-ATP carrier protein 5 (ADP/ATP translocase 5).  
 GN TLCE OR TLC5 OR RP739.  
 OS Rickettsia prowazekii.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;  
 OC Rickettsiaceae; Rickettsiae; Rickettsia.  
 OX NCBI\_TaxID=782;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Madrid E;  
 RX MEDLINE=99039499; PubMed=9823893;  
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,  
 RA Scharitz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,  
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;



RT "The genome sequence of Rickettsia prowazekii and the origin of  
 RL mitochondria.";  
 RL Nature 396:133-140(1998).  
 RN [2]  
 RP SEQUENCE OF 325-500 FROM N.A.  
 RC STRAIN=Madrid E;  
 RX MEDLINE=97419517; PubMed=9274032;  
 RA Andersson J.C., Andersson S.G.E.;  
 RT "Genomic rearrangements during evolution of the obligate  
 RT intracellular parasite Rickettsia prowazekii as inferred from an  
 RT analysis of 52015 bp nucleotide sequence.";  
 RL Microbiology 143:2783-2795(1997).  
 CC -!- FUNCTION: PROVIDES THE RICKETTSIAL CELL WITH HOST ATP IN  
 CC EXCHANGE FOR RICKETTSIAL ADP. THIS IS AN OBLIGATE EXCHANGE SYSTEM.  
 CC THIS ENERGY ACQUIRING ACTIVITY IS AN IMPORTANT COMPONENT OF  
 CC RICKETTSIAL PARASITISM (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: BELONGS TO THE ADP/ATP TRANSLOCASE TLC FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC  
 CC EMBL; AJ235273; CAA15167.1; --  
 CC EMBL; Y11778; CAA72457.1; --  
 DR InterPro; IPR004667; ADP\_ATP\_car.  
 DR Pfam; PF03219; TLC; 1.  
 DR TIGRFAMs; TIGR00769; AAA; 1.  
 KW Transmembrane; Transport; ATP-binding; Multigene family;  
 KW Complete proteome.  
 FT TRANSMEM 26 46 POTENTIAL.  
 FT TRANSMEM 62 82 POTENTIAL.  
 FT TRANSMEM 94 114 POTENTIAL.  
 FT TRANSMEM 149 169 POTENTIAL.  
 FT TRANSMEM 184 204 POTENTIAL.  
 FT TRANSMEM 224 244 POTENTIAL.  
 FT TRANSMEM 287 307 POTENTIAL.  
 FT TRANSMEM 328 348 POTENTIAL.  
 FT TRANSMEM 357 377 POTENTIAL.  
 FT TRANSMEM 381 401 POTENTIAL.  
 FT TRANSMEM 469 489 POTENTIAL.  
 SQ SEQUENCE 500 AA; 57073 MW; FE3DB48D08CF5F72 CRC64;  
 Query Match 52.9%; Score 37; DB 1; Length 500;  
 Best Local Similarity 54.5%; Pred. No. 25;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 CFAWKENMKV 11  
 Db 482 CFAWIVAVRKI 492  
 RESULT 14  
 ID VP4\_ROTGA STANDARD; PRT; 749 AA.  
 AC Q04916;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE Outer capsid protein VP4 (Hemagglutinin) (Outer layer protein VP4).  
 GN S4.  
 OS Rotavirus (group B / strain ADRV) (Adult diarrhea rotavirus).  
 OC Viruses; dsRNA viruses; Reoviridae; Rotavirus.  
 OC NCBI\_TaxID=12705;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93233240; PubMed=8386274;  
 RA Mackow E.R., Werner-Eckert R., Fay M.E., Tao H., Chen G.-M.;  
 RT "Identification and baculovirus expression of the VP4 protein of the

RT human group B rotavirus ADRV.";  
 RL J. Virol. 67:2730-2738(1993).  
 CC -!- SUBCELLULAR LOCATION: Outer capsid.  
 CC -!- SIMILARITY: BELONGS TO THE ROTAVIRUS VP4 PROTEIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC  
 CC EMBL; M91434; AAA47338.1; --  
 DR InterPro; IPR000416; Cap\_VP4.  
 DR Pfam; PF00426; VP4; 1.  
 KW Coat protein; Glycoprotein.  
 FT CARBOHYD 29 29  
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 109 109 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 133 133 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 407 407 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 527 527 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 568 568 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 620 620 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 681 681 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 698 698 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 749 AA; 84362 MW; D123527DEAE0F21 CRC64;  
 Query Match 52.9%; Score 37; DB 1; Length 749;  
 Best Local Similarity 50.0%; Pred. No. 38;  
 Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 CFAWKENMKV 12  
 Db 195 CFTMDMNCANVR 206  
 RESULT 15  
 ID WEST\_SALTI STANDARD; PRT; 116 AA.  
 AC P04336;  
 DT 20-MAR-1987 (Rel. 04, Created)  
 DT 20-MAR-1987 (Rel. 04, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Mercuroic transport protein (Mercury ion transport protein).  
 GN MERT OR HCM1\_234C.  
 OS Salmonella typhi; and  
 OS Shigella flexneri.  
 OG Plasmid pHCMI, and Plasmid IncFII NRI.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Salmonella.  
 OC NCBI\_TaxID=601, 623;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S.typhi; STRAIN=CT18; PLASMID=pHCMI;  
 RX MEDLINE=21534947; PubMed=11677608;  
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,  
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,  
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,  
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,  
 RA Krögh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,  
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,  
 RA Whitehead S., Barrell B.G.;  
 RT "Complete genome sequence of a multiple drug resistant Salmonella  
 RT enterica serovar Typhi CT18.";  
 RL Nature 413:848-852(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S.flexneri;  
 RX MEDLINE=85014891; PubMed=6091128;  
 RA Misra T.K., Brown N.L., Fritzinger D.C., Pridmore R.D., Barnes W.M.,

```

RA Haberstroh L., Silver S.;
RT "Mercuric ion-resistance operons of plasmid R100 and transposon
RT Tn501: the beginning of the operon including the regulatory region
RT and the first two structural genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:5975-5979 (1984).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; PLASMID=IncFII NR1; TRANSPOSON=Tn21;
RX MEDLINE=85159407; PubMed=6530603;
RA Barthelemy P., Gilbert P., Jackson W.J., Jones C.S., Summers A.O.,
RA Wisdom S.;
RT "The DNA sequence of the mercury resistance operon of the IncFII
RT plasmid NR1.";
RL J. Mol. Appl. Genet. 2:601-619 (1984).
CC -1- FUNCTION: INVOLVED IN MERCURIC TRANSPORT. PASSES A HG(2+) ION
CC FROM THE PERIPLASMIC MERP PROTEIN TO THE MERCURIC REDUCTASE
CC (NERA)
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL513383; CAD09816.1; -
DR EMBL; J01730; AAA92261.1; -
DR EMBL; K03089; AAB53075.1; -
DR PIR; A04458; QJEBHT.
DR InterPro; IPR003457; Transprt_MerT.
DR Pfam; PF02411; MerT; 1.
DR Transmembrane; Mercuric resistance; Inner membrane; Mercury; Plasmid;
KW Transport; Mercuric resistance; Inner membrane; Mercury; Plasmid;
KW Transmembrane; Transportable element; Complete proteome.
FT TRANSMEM 16 36 POTENTIAL.
FT TRANSMEM 46 66 POTENTIAL.
FT TRANSMEM 94 114 POTENTIAL.
FT METAL 24 24 HG(2+) (POTENTIAL).
FT METAL 25 25 HG(2+) (POTENTIAL).
FT METAL 76 76 HG(2+) (POTENTIAL).
FT METAL 82 82 HG(2+) (POTENTIAL).
SQ SEQUENCE 116 AA; 12521 MW; 8CF0744F4B0F6EE4 CRC64;

```

```

Query Match      51.4%; Score 36; DB 1; Length 116;
Best Local Similarity 63.6%; Pred. No. 8.3;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 2 FAWKKNMKVR 12
   |||||
Db 64 FAWKKIYRPVQ 74

```

```

Search completed: February 21, 2003, 07:51:41
Job time : 6.2093 secs

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:44:43 ; Search time 21.6744 Seconds  
(without alignments)  
114.078 Million cell updates/sec

Title: US-09-743-107B-94

Perfect score: 70

Sequence: 1 CFAWKNRKR 12

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriaph:\*
- 17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	80.0	711	4 Q8TCD2	Q8tcd2 homo sapien
2	52	74.3	38	4 Q9UCV5	Q9ucv5 homo sapien
3	46	65.7	511	16 Q82462	Q82462 salmonella
4	43	61.4	393	16 Q25611	Q25611 helicobacte
5	41	58.6	33	6 Q9TR80	Q9tr80 ovis aries
6	41	58.6	550	11 Q9JUZ5	Q9jjz5 mus musculu
7	41	58.6	584	17 Q9HPA3	Q9hpa3 halobacteri
8	40	57.1	62	16 Q8R9U1	Q8r9u1 thermococae
9	40	57.1	206	4 Q9BXW1	Q9bxw1 homo sapien
10	40	57.1	301	4 Q8TDU2	Q8tdu2 homo sapien
11	40	57.1	374	4 Q60735	Q60735 homo sapien
12	40	57.1	374	4 Q96XU8	Q96xu8 homo sapien
13	40	57.1	374	11 Q9JUX1	Q9jux1 mus musculu
14	40	57.1	386	16 Q9PQS6	Q9pq6 ureaplasma
15	39	55.7	114	11 Q63104	Q63104 rattus norv
16	39	55.7	121	16 Q8U6K3	Q8u6k3 agrobacteri

17	39	55.7	205	16 Q986A0	Q986a0 rhizobium 1
18	39	55.7	206	16 Q8UHC2	Q8uhc2 agrobacteri
19	39	55.7	206	16 Q92RH8	Q92rh8 rhizobium m
20	39	55.7	208	16 Q8YFK3	Q8yfk3 bruceella me
21	39	55.7	234	11 Q63112	Q63112 rattus norv
22	39	55.7	275	5 Q93780	Q93780 caenorhabdi
23	39	55.7	282	16 Q98Q19	Q98q19 mycoplasma
24	39	55.7	341	11 Q8R2A4	Q8r2a4 mus musculu
25	39	55.7	393	16 Q9ZKP4	Q9zkp4 helicobacte
26	39	55.7	447	5 Q17549	Q17549 caenorhabdi
27	39	55.7	519	10 Q93833	Q93833 arabidopsis
28	39	55.7	553	4 Q9NZL7	Q9nzt7 homo sapien
29	39	55.7	554	4 Q9NY67	Q9ny67 homo sapien
30	39	55.7	558	4 Q9UJK6	Q9ufk6 homo sapien
31	39	55.7	720	10 Q9CA22	Q9ca22 arabidopsis
32	38	54.3	81	15 Q90863	Q90863 human immun
33	38	54.3	205	8 Q98RR2	Q98rr2 guillardia
34	38	54.3	273	2 Q31090	Q31090 rhizobium 1
35	38	54.3	318	2 Q9ADZ8	Q9adz8 agrobacteri
36	38	54.3	886	5 Q9VZV1	Q9vzv1 drosophila
37	37	52.9	196	4 Q14557	Q14557 homo sapien
38	37	52.9	225	5 Q21197	Q21197 caenorhabdi
39	37	52.9	232	4 Q98H38	Q98h38 homo sapien
40	37	52.9	236	10 Q9SK05	Q9sk05 arabidopsis
41	37	52.9	246	4 Q96GY3	Q96gy3 homo sapien
42	37	52.9	289	10 Q9C6N2	Q9c6n2 arabidopsis
43	37	52.9	302	16 Q987S1	Q987s1 rhizobium 1
44	37	52.9	317	5 Q44841	Q44841 caenorhabdi
45	37	52.9	502	10 Q85237	Q85237 arabidopsis

## ALIGNMENTS

### RESULT 1

Q8TCD2 ID Q8TCD2 PRELIMINARY; PRT; 711 AA.  
AC Q8TCD2;  
DT 01-JUN-2002 (Tremblrel. 21, Created)  
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE Lactotransferrin.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN NCBI\_TaxID=9606;  
[1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PROSTATE;  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC022347; AAH22347.1;  
SQ SEQUENCE 711 AA; 78327 MW; 1B9C7BE097C45FAF CRC64;

Query Match 80.0%; Score 56; DB 4; Length 711;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFAWKNRKR 11

Db 39 CFQWQNRKR 49

### RESULT 2

Q9UCV5 ID Q9UCV5 PRELIMINARY; PRT; 38 AA.  
AC Q9UCV5;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE Lactoferrin homolog (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96081613; PubMed=8551595;  
RA Sato I.;  
RT "Characterization of the 84-kDa protein with ABH activity in human  
semenal plasma";  
RL Jpn. J. Legal Med. 49:281-293(1995).  
RA HSP; P02788; 1BK.  
DR InterPro: IPR001156; Transferrin.  
DR Pfam: PF00405; transferrin; 1.  
SQ SEQUENCE 38 AA; 4459 MW; 0402F490B5EBDBEB CRC64;  
  
Query Match 74.3%; Score 52; DB 4; Length 38;  
Best Local Similarity 81.8%; Pred. No. 0.026;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 FAWKNNMKVR 12  
|:|:|:|:|:|  
Db 21 FQWNNMKVR 31  
|:|:|:|:|:|  
  
RESULT 3  
O82462 PRELIMINARY; PRT; 511 AA.  
ID Q82462  
AC Q82462  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DE Hypothetical protein STV3070.  
GN STV3070  
OS Salmonella typhi.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Salmonella.  
OX NCBI\_TaxID=601;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21534947; PubMed=11677608;  
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,  
Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,  
Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,  
Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
Faulstich T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,  
Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,  
Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,  
Whitehead S., Barrall B.G.;  
RT "Complete genome sequence of a multiple drug resistant Salmonella  
enterica serovar Typhi CT18.";  
RL Nature 413:848-852(2001).  
DR EMBL: AL627276; CAD06049.1; --  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 511 AA; 58126 MW; E2DD124E10D178B CRC64;  
  
Query Match 65.7%; Score 46; DB 16; Length 511;  
Best Local Similarity 66.7%; Pred. No. 4.4;  
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 CFAWNNMKVR 12  
|:|:|:|:|:|  
Db 350 CFAWNNMKVR 361  
|:|:|:|:|:|  
  
RESULT 4  
O25611 PRELIMINARY; PRT; 393 AA.  
ID O25611  
AC O25611  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DE 3-deoxy-D-manno-octulosonic-acid transferase (KOTA).  
GN HP0957.

OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=26695 / ATCC 700392;  
RA MEDLINE=97394467; PubMed=9252185;  
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,  
Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,  
Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,  
Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,  
McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,  
Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,  
Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watney L., Wallin E.,  
Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,  
Venter J.C.;  
RT "The complete genome sequence of the gastric pathogen Helicobacter  
pylori";  
RL Nature 388:539-547(1997).  
DR EMBL: AE000604; AAD08000.1; --  
DR TIGR: HP0957; --  
KW Hypothetical protein; Transferrin; Complete proteome.  
SQ SEQUENCE 393 AA; 45622 MW; 6428BA332154F46 CRC64;  
  
Query Match 61.4%; Score 43; DB 16; Length 393;  
Best Local Similarity 45.5%; Pred. No. 12;  
Matches 5; Conservative 6; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 FAWKNNMKVR 12  
|:|:|:|:|:|  
Db 110 FAWKNNMKVR 120  
|:|:|:|:|:|  
  
RESULT 5  
Q9TR80 PRELIMINARY; PRT; 33 AA.  
ID Q9TR80  
AC Q9TR80  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DE Lactoferrin (Fragment).  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Ruminantia; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=95127729; PubMed=7827104;  
RA Qian Z.Y., Jolles P., Migliore-Samou D., Fiat A.M.;  
RL Biochim. Biophys. Acta 1243:25-32(1995).  
DR HSP; O7698; 1CB2.  
DR InterPro: IPR001156; Transferrin.  
DR Pfam: PF00405; transferrin; 1.  
SQ SEQUENCE 33 AA; 3914 MW; D1904CB15A73961 CRC64;  
  
Query Match 58.6%; Score 41; DB 6; Length 33;  
Best Local Similarity 45.5%; Pred. No. 2;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 CFAWNNMKVR 11  
|:|:|:|:|:|  
Db 19 CYQWNNMKVR 29  
|:|:|:|:|:|  
  
RESULT 6  
Q9JUZ5 PRELIMINARY; PRT; 550 AA.  
ID Q9JUZ5  
AC Q9JUZ5  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DE 3-deoxy-D-manno-octulosonic-acid transferase (KOTA).  
GN HP0957.

DE Hypothetical 61.5 kDa protein.  
GN EGF\_L6 OR W80.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Franco B.,  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20241927; PubMed=10777661;  
RA Buchner G., Orfanelli U., Quaderi N., Bassi M.T., Andolfi G.;  
RT "Identification of a new EGF-repeat-containing gene from human Xp22:  
RT A candidate for developmental disorders";  
RL Genomics 65:16-23(2000).  
CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.  
DR EMBL; AJ245672; CAB92138.1; -.  
DR HSP; P35555; 1EMN.  
DR MGD; MGI:1858599; Egfl6.  
DR InterPro; IPR001152; Asx\_hydroxyl.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR000998; MAM\_domain.  
DR Pfam; PF00008; EGF; 4.  
DR Pfam; PF00629; MAM; 1.  
DR SMART; SM00179; EGF\_CA; 3.  
DR SMART; SM00001; EGF\_Like; 2.  
DR SMART; SM00137; MAM; 1.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 3.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE; PS01186; EGF\_2; 2.  
DR PROSITE; PS01187; EGF\_CA; 3.  
DR PROSITE; PS00060; MAM\_2; 1.  
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation;  
KW Hypothetical protein; Repeat.  
SQ SEQUENCE 550 AA; 61520 MW; DEF936325C9F31B3 CRC64;  
Query Match 58.6%; Score 41; DB 11; Length 550;  
Best Local Similarity 56.5%; Pred. No. 37;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 1 CFARKNNRKV 11  
Db 48 CYGWRNNKGV 58  
RESULT 7  
Q9HPA3 PRELIMINARY; PRT; 594 AA.  
AC Q9HPA3;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Vngl732c.  
GN VNGI732C.  
OS Halobacterium sp. (strain NRC-1).  
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
OC Halobacteriaceae; Halobacterium.  
OX NCBI\_TaxID=64091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20504483; PubMed=11016950;  
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,  
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,  
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,  
RA Leithauer S., Keller K., Cruz R., Danson M.J., Hough D.W.,  
RA Maddocks D.G., Jablonksi P.E., Krebs M.P., Angevine C.M., Dale H.,  
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,  
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;  
RT "Genome sequence of Halobacterium species NRC-1";

RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
DR EMBL; AB005078; AAG19967.1; -.  
DR InterPro; IPR001646; Septide\_repeat.  
DR InterPro; IPR001622; K-channel\_pore.  
DR Pfam; PF00805; Pentapeptide\_2.  
KW Complete proteome.  
SQ SEQUENCE 584 AA; 65151 MW; 21BF5DF0486CCC6 CRC64;  
Query Match 58.6%; Score 41; DB 17; Length 584;  
Best Local Similarity 41.7%; Pred. No. 39;  
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
QY 1 CFARKNNRKV 12  
Db 445 CFTWRKONERR 456  
RESULT 8  
Q8R9U1 PRELIMINARY; PRT; 62 AA.  
AC Q8R9U1;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Ribosomal protein L28.  
GN RMB OR TTE1495.  
OS Thermoanaerobacter tengcongensis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
OC Thermoanaerobacteriales; Thermoanaerobacteriaceae; Thermoanaerobacter.  
OX NCBI\_TaxID=119072;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MB4T / JCM11007;  
RX MEDLINE=21992816; PubMed=11997336;  
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
RA Tan H., Chen R., Wang J., Yu J., Yang H.;  
RT "A complete sequence of T. tengcongensis genome";  
RL Genome Res. 12:689-700(2002).  
DR EMBL; AE013107; AAM24713.1; -.  
KW Complete proteome.  
SQ SEQUENCE 62 AA; 7037 MW; CAADB605C81D495D CRC64;  
Query Match 57.1%; Score 40; DB 16; Length 62;  
Best Local Similarity 77.8%; Pred. No. 5.9;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 4 WKNNRKV 12  
Db 28 WKNRKV 36  
RESULT 9  
Q9BXW1 PRELIMINARY; PRT; 206 AA.  
AC Q9BXW1;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE PNAS-125.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX Yu W.-Q., Sun B.-Z., Chai Y.-B., Zhu F., Liu X.-S., Li Z., Lu F.,  
RA Yan W., Yang H., Zhao Z.-L.;  
RT "Human acute promyelocytic leukemia cell line NB4's  
RT apoptosis/differentiation related genes";  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF271183; AAK07542.1; -.  
DR InterPro; IPR000717; PCI.

DR Pfam; PF01399; PCI; 1.  
DR SMART; SM00088; PINT; 1.  
SQ SEQUENCE 206 AA; 23754 MW; 940D5B66271A44B4 CRC64;

Query Match 57.1%; Score 40; DB 4; Length 206;  
Best Local Similarity 60.0%; Pred. No. 20;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 AKRNNMKVR 12  
DB 172 AKQNLNKVK 181  
||||:|:

## RESULT 10

ID Q8TDU2 PRELIMINARY; PRT; 301 AA.  
AC Q8TDU2;  
DT 01-JUN-2002 (TRENBLrel. 21, Created)  
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
DE Putative G-protein coupled receptor.  
GN GPCR.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Takeda S., Kadowaki S., Haga T., Takesu H., Mitaku S.;  
RT "Identification of G protein-coupled receptor genes from the human  
genome sequence."  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
RW EMBL; AB083609; BAB89322.1; -.  
KW Receptor.  
SQ SEQUENCE 301 AA; 33557 MW; AB7F9792957BFCA6 CRC64;

Query Match 57.1%; Score 40; DB 4; Length 301;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CFAWKR 6  
DB 296 CFAWKR 301  
|||||

## RESULT 11

ID O60735 PRELIMINARY; PRT; 374 AA.  
AC O60735;  
DT 01-AUG-1998 (TRENBLrel. 07, Created)  
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)  
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
DE Gal7 protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BLOOD;  
RA Zhao Z., Huang X., Li N., Zhu X., Cao X.;  
RT "A novel gene from human dendritic cell."  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
RW EMBL; AF064603; AAC17108.1; -.  
DR InterPro; IPR000717; PCI.  
DR Pfam; PF01399; PCI; 1.  
DR SMART; SM00088; PINT; 1.  
SQ SEQUENCE 374 AA; 42512 MW; 2CFF2BCB70F997A7 CRC64;

Query Match 57.1%; Score 40; DB 4; Length 374;  
Best Local Similarity 60.0%; Pred. No. 38;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 AKRNNMKVR 12  
DB 356 AKQNLNKVK 365  
||||:|:

## RESULT 12

ID Q96KM8 PRELIMINARY; PRT; 374 AA.  
AC Q96KM8;  
DT 01-DEC-2001 (TRENBLrel. 19, Created)  
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
DE DJ69B10.1 (Gal7 protein) (Hypothetical 42.5 kDa protein).  
GN DJ69B10.1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Cobley V.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=MUSCLE;  
RA Strausberg R.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL078477; CAC88392.1; -.  
DR EMBL; BC019103; AAH19103.1; -.  
DR InterPro; IPR000717; PCI.  
DR Pfam; PF01399; PCI; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 374 AA; 42503 MW; 63736CA2B093D794 CRC64;

Query Match 57.1%; Score 40; DB 4; Length 374;  
Best Local Similarity 60.0%; Pred. No. 38;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 AKRNNMKVR 12  
DB 356 AKQNLNKVK 365  
||||:|:

## RESULT 13

ID Q99JX4 PRELIMINARY; PRT; 374 AA.  
AC Q99JX4;  
DT 01-JUN-2001 (TRENBLrel. 17, Created)  
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
DE Similar to dendritic cell protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC005598; AAH05598.1; -.  
DR InterPro; IPR000717; PCI.  
DR Pfam; PF01399; PCI; 1.  
DR SMART; SM00088; PINT; 1.  
SQ SEQUENCE 374 AA; 42517 MW; 09845D87E792B783 CRC64;

Query Match 57.1%; Score 40; DB 11; Length 374;  
Best Local Similarity 60.0%; Pred. No. 38;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 AKRNNMKVR 12  
DB 356 AKQNLNKVK 365  
||||:|:

Qy 2 FAWKRNMRKV 11  
 ||| | :||:  
 Db 42 FAWYRGLRKI 51

Search completed: February 21, 2003, 08:00:47  
 Job time : 22.6744 secs

RESULT 14  
 Q9POS6  
 ID Q9POS6 PRELIMINARY; PRT; 386 AA.  
 AC Q9POS6;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
 DE Serine/threonine kinase.  
 GN FKN OR UU216.  
 OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 OC Mycoplasmataceae; Ureaplasma.  
 OX NCBI\_TaxID=134821;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SEROVAR 3;  
 RX MEDLINE=20500219; PubMed=11048724;  
 RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,  
 RA Casell G.H.;  
 RT "The complete sequence of the mucosal pathogen Ureaplasma  
 RT urealyticum.";  
 RL Nature 407:757-762(2000).  
 DR EMBL; AE002120; AAF30624.1; -;  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; UNKNOWN 1.  
 KW ATP-binding; Transferase; Complete proteome.  
 SQ SEQUENCE 386 AA; 45241 MW; 789374CC15D1CAD9 CRC64;

Query Match 57.1%; Score 40; DB 16; Length 386;  
 Best Local Similarity 54.5%; Pred.No. 39;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CFARKNMRKV 11  
 ||| | :||:  
 Db 290 CFARKEDNKL 300

RESULT 15  
 Q63104  
 ID Q63104 PRELIMINARY; PRT; 114 AA.  
 AC Q63104;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-JAN-1999 (TReMBLrel. 09, Last sequence update)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
 DE Carcinoembryonic antigen (fragment).  
 GN CEA4.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BD II; TISSUE=LIVER;  
 RX MEDLINE=89214106; PubMed=2708349;  
 RA Kodellja V., Lucas K., Barnert S., von Kleist S., Thompson J.A.,  
 RA Zimmermann W.A.;  
 RT "Identification of a carcinoembryonic antigen gene family in the rat:  
 RT Analysis of the N-terminal domains reveals immunoglobulin-like,  
 RT hypervariable regions.";  
 RL J. Biol. Chem. 264:6906-6912(1989).  
 DR EMBL; M60026; AAA40911.1; -;  
 FT NON TER 1  
 SQ SEQUENCE 114 AA; 12832 MW; 3AE108689B061686 CRC64;

Query Match 55.7%; Score 39; DB 11; Length 114;  
 Best Local Similarity 50.0%; Pred.No. 17;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:37:21 ; Search time 28.093 Seconds  
(without alignments)  
56.918 Million cell updates/sec

Title: US-09-743-107B-95  
Perfect score: 58  
Sequence: 1 CFAWQRAWRKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
1: /SID32/gcgdata/geneseq/genesecp-emb1/AA1980.DAT.\*  
2: /SID32/gcgdata/geneseq/genesecp-emb1/AA1981.DAT.\*  
3: /SID32/gcgdata/geneseq/genesecp-emb1/AA1982.DAT.\*  
4: /SID32/gcgdata/geneseq/genesecp-emb1/AA1983.DAT.\*  
5: /SID32/gcgdata/geneseq/genesecp-emb1/AA1984.DAT.\*  
6: /SID32/gcgdata/geneseq/genesecp-emb1/AA1985.DAT.\*  
7: /SID32/gcgdata/geneseq/genesecp-emb1/AA1986.DAT.\*  
8: /SID32/gcgdata/geneseq/genesecp-emb1/AA1987.DAT.\*  
9: /SID32/gcgdata/geneseq/genesecp-emb1/AA1988.DAT.\*  
10: /SID32/gcgdata/geneseq/genesecp-emb1/AA1989.DAT.\*  
11: /SID32/gcgdata/geneseq/genesecp-emb1/AA1990.DAT.\*  
12: /SID32/gcgdata/geneseq/genesecp-emb1/AA1991.DAT.\*  
13: /SID32/gcgdata/geneseq/genesecp-emb1/AA1992.DAT.\*  
14: /SID32/gcgdata/geneseq/genesecp-emb1/AA1993.DAT.\*  
15: /SID32/gcgdata/geneseq/genesecp-emb1/AA1994.DAT.\*  
16: /SID32/gcgdata/geneseq/genesecp-emb1/AA1995.DAT.\*  
17: /SID32/gcgdata/geneseq/genesecp-emb1/AA1996.DAT.\*  
18: /SID32/gcgdata/geneseq/genesecp-emb1/AA1997.DAT.\*  
19: /SID32/gcgdata/geneseq/genesecp-emb1/AA1998.DAT.\*  
20: /SID32/gcgdata/geneseq/genesecp-emb1/AA2000.DAT.\*  
21: /SID32/gcgdata/geneseq/genesecp-emb1/AA2001.DAT.\*  
22: /SID32/gcgdata/geneseq/genesecp-emb1/AA2002.DAT.\*  
23: /SID32/gcgdata/geneseq/genesecp-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	100.0	12	21	AA198095
2	63	92.6	12	21	AA198074
3	62	91.2	12	21	AA198070
4	59	85.8	12	21	AA198089
5	59	86.8	12	21	AA198090
6	59	86.8	12	21	AA198093
7	58	85.3	12	21	AA198086
8	58	85.3	12	21	AA198094
9	57	83.8	12	21	AA198038
10	57	83.8	12	21	AA198046

11	57	83.8	12	21	AA198047	Human lactoferrin
12	57	83.8	13	21	AA198037	Human lactoferrin
13	57	83.8	13	21	AA198048	Human lactoferrin
14	57	83.8	13	21	AA198049	Human lactoferrin
15	57	83.8	14	21	AA198036	Human lactoferrin
16	57	83.8	14	21	AA198050	Human lactoferrin
17	57	83.8	14	21	AA198051	Human lactoferrin
18	57	83.8	15	17	AA198554	Peptide for anti-u
19	57	83.8	15	21	AA198035	Human lactoferrin
20	57	83.8	15	21	AA198062	Human lactoferrin
21	57	83.8	15	21	AA198063	Human lactoferrin
22	57	83.8	16	21	AA198031	Human lactoferrin
23	57	83.8	16	21	AA198064	Human lactoferrin
24	57	83.8	16	21	AA198065	Human lactoferrin
25	57	83.8	17	21	AA198034	Human lactoferrin
26	57	83.8	17	21	AA198066	Human lactoferrin
27	57	83.8	17	21	AA198067	Human lactoferrin
28	57	83.8	18	15	AA198352	Human lactoferrin
29	57	83.8	18	17	AA198397	Advanced glycosyla
30	57	83.8	18	21	AA198033	Human lactoferrin
31	57	83.8	19	21	AA198867	Amino acid sequenc
32	57	83.8	19	21	AA198032	Human lactoferrin
33	57	83.8	20	13	AA1981810	Anti microbial pep
34	57	83.8	20	14	AA1984841	Lactoferrin-relate
35	57	83.8	20	15	AA198530	Lactoferrin derive
36	57	83.8	20	15	AA198531	Lactoferrin derive
37	57	83.8	20	15	AA1987461	Lactoferrin derive
38	57	83.8	20	15	AA1987462	Lactoferrin derive
39	57	83.8	20	16	AA1984698	Bovine lactoferrin
40	57	83.8	20	16	AA1984699	Bovine lactoferrin
41	57	83.8	20	16	AA1980263	Anti-parasitic lac
42	57	83.8	20	16	AA1980264	Anti-parasitic lac
43	57	83.8	20	17	AA1988553	Peptide for anti-u
44	57	83.8	20	17	AA1981852	Lactoferrin-derive
45	57	83.8	20	17	AA1980345	Lactoferrin-derive

#### ALIGNMENTS

RESULT 1  
AA198095  
ID AA198095 standard; Peptide; 12 AA.  
XX AA198095;  
AC AA198095;  
XX AA198095;  
DT 25-APR-2000 (first entry)  
XX Human lactoferrin derived peptide SEQ ID NO:95.  
DE Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX Homo sapiens.  
OS Synthetic.  
XX WO200001730-A1.  
XX 13-JAN-2000.  
XX 06-JUL-1999; 99WO-S801230.  
XX 06-JUL-1998; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
XX (ASCI-) A+ SCI INVEST AB.  
PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX WPI; 2000-147388/13.



XX New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
XX Claim 22; Page 38; 102pp; English.  
XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.  
XX Sequence 12 AA;  
SQ Query Match 100.0%; Score 68; DB 21; Length 12;  
Best Local Similarity 100.0%; Pred. No. 5.6e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CFAWQRAWRKVR 12  
Db |||||  
1 CFAWQRAWRKVR 12  
RESULT 2  
AAY78074  
ID AAY78074 standard; Peptide; 12 AA.  
AC AAY78074;  
XX 25-APR-2000 (first entry)  
DT Human lactoferrin derived peptide SEQ ID NO:74.  
DE Human; lactoferrin; modification; infection; inflammation; tumour;  
XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX Homo sapiens.  
OS Synthetic.  
XX WO200001730-A1.  
PN 13-JAN-2000.  
PD 06-JUL-1999; 99WO-SE01230.  
PF 06-JUL-1998; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
XX (ASCI-) A+ SCI INVEST AB.  
PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
PI WPI; 2000-147388/13.  
XX New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
XX Claim 22; Page 35; 102pp; English.  
XX AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.  
XX Sequence 12 AA;  
SQ Query Match 92.6%; Score 63; DB 21; Length 12;  
Best Local Similarity 91.7%; Pred. No. 0.00038;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CFAWQRAWRKVR 12  
Db |||||  
1 CFAWQRAWRKVR 12  
RESULT 3  
AAY78070  
ID AAY78070 standard; Peptide; 12 AA.  
AC AAY78070;  
XX 25-APR-2000 (first entry)  
DT Human lactoferrin derived peptide SEQ ID NO:70.  
DE Human; lactoferrin; modification; infection; inflammation; tumour;  
XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX Homo sapiens.  
OS Synthetic.  
XX WO200001730-A1.  
PN 13-JAN-2000.  
PD 06-JUL-1999; 99WO-SE01230.  
PF 06-JUL-1998; 98SE-0002441.  
PR 17-JUL-1998; 98SE-0002562.  
PR 29-DEC-1998; 98SE-0004614.  
XX (ASCI-) A+ SCI INVEST AB.  
PA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
PI WPI; 2000-147388/13.  
XX New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
XX Claim 22; Page 35; 102pp; English.  
XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides can also be used  
CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;

Query Match 91.2%; Score 62; DB 21; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 0.00057;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CFQWQRMKVR 12  
 Db 1 CFQWQRMKVR 12

RESULT 4

AAV78089  
 ID AAY78089 standard; Peptide; 12 AA.

XX AC AAY78089;

XX DT 25-APR-2000 (first entry)

DE Human lactoferrin derived peptide SEQ ID NO:89.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX OS Homo sapiens.  
 OS Synthetic.

XX PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

XX PS Claim 22; Page 37; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;

Query Match 86.8%; Score 59; DB 21; Length 12;  
 Best Local Similarity 83.3%; Pred. No. 0.0018;  
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 CFQWQRMKVR 12  
 Db 1 CFQWQRMKVR 12

RESULT 5

AAV78090  
 ID AAY78090 standard; Peptide; 12 AA.

XX AC AAY78090;

XX DT 25-APR-2000 (first entry)

DE Human lactoferrin derived peptide SEQ ID NO:90.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX OS Homo sapiens.  
 OS Synthetic.

XX PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

XX PS Claim 22; Page 37; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;

Query Match 86.8%; Score 59; DB 21; Length 12;  
 Best Local Similarity 83.3%; Pred. No. 0.0018;  
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFAWQRAVRKVR 12  
 XX |||||  
 AC 1 CFQWQRAVRKVR 12  
 DB |||||

## RESULT 6

AAV78093  
 ID AAV78093 standard; Peptide; 12 AA.

XX AAV78093;  
 AC  
 DT 25-APR-2000 (first entry)  
 XX

DE Human lactoferrin derived peptide SEQ ID NO:93.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

OS Homo sapiens.  
 OS Synthetic.

XX WO200001730-A1.

PN 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCII-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

PI WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections.

PT Inflammations and tumors and for use in infant formula food -

XX Claim 22; Page 38; 102pp; English.

XX AAV78001 to AAV78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 12 AA;

Query Match 86.8%; Score 59; DB 21; Length 12;

Best Local Similarity 83.3%; Pred. No. 0.0018;

Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFAWQRAVRKVR 12

DB |||||

## RESULT 7

AAV78086

ID AAV78086 standard; Peptide; 12 AA.

XX AAV78086;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:86.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.

OS Synthetic.

XX WO200001730-A1.

PN 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCII-) A+ SCI INVEST AB.

XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

PI WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections.

PT Inflammations and tumors and for use in infant formula food -

XX Claim 22; Page 36; 102pp; English.

XX AAV78001 to AAV78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 12 AA;

Query Match 85.3%; Score 58; DB 21; Length 12;

Best Local Similarity 83.3%; Pred. No. 0.0027;

Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFAWQRAVRKVR 12

DB |||||

## RESULT 8

AAV78094

ID AAV78094 standard; Peptide; 12 AA.

XX AAV78094;

XX 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:94.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200001730-A1.  
 XX  
 PD 13-JAN-2000.  
 XX  
 PF 06-JUL-1999; 99WO-SE01230.  
 XX  
 PR 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 PA (ASCI-) A+ SCI INVEST AB.  
 XX  
 PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX  
 DR WPI; 2000-147388/13.  
 XX  
 PT New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX  
 PS Claim 22; Page 38; 102pp; English.  
 XX  
 CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 SQ Sequence 12 AA;  
 Query Match 85.3%; Score 58; DB 21; Length 12;  
 Best Local Similarity 83.3%; Pred. No. 0.0027;  
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFAMQRAMRKVR 12  
 |||||  
 DB 1 CFAMQNNRKVR 12  
 |||||  
 RESULT 9  
 AAY78038  
 ID AAY78038 standard; Peptide; 12 AA.  
 XX  
 AC AAY78038;  
 XX  
 DT 25-APR-2000 (first entry)  
 XX  
 DE Human lactoferrin derived peptide SEQ ID NO:38.  
 XX  
 KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.

OS Synthetic.  
 XX  
 PN WO200001730-A1.  
 XX  
 PD 13-JAN-2000.  
 XX  
 PF 06-JUL-1999; 99WO-SE01230.  
 XX  
 PR 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 PA (ASCI-) A+ SCI INVEST AB.  
 XX  
 PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX  
 DR WPI; 2000-147388/13.  
 XX  
 PT New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX  
 PS Claim 12; Page 70; 102pp; English.  
 XX  
 CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 SQ Sequence 12 AA;  
 Query Match 83.8%; Score 57; DB 21; Length 12;  
 Best Local Similarity 83.3%; Pred. No. 0.0039;  
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CFAMQRAMRKVR 12  
 |||||  
 DB 1 CFAMQNNRKVR 12  
 |||||  
 RESULT 10  
 AAY78046  
 ID AAY78046 standard; Peptide; 12 AA.  
 XX  
 AC AAY78046;  
 XX  
 DT 25-APR-2000 (first entry)  
 XX  
 DE Human lactoferrin derived peptide SEQ ID NO:46.  
 XX  
 KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200001730-A1.  
 XX  
 PD 13-JAN-2000.  
 XX  
 PF 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 15; Page 35; 102pp; English.  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX SQ Sequence 12 AA;  
 Query Match 83.8%; Score 57; DB 21; Length 12;  
 Best Local Similarity 83.3%; Pred. No. 0.0039;  
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CFQWQRNMRKVR 12  
 Db |||||  
 1 CFQWQRNMRKVR 12  
 RESULT 11  
 AAY78047  
 ID AAY78047 standard; Peptide; 12 AA.  
 AC AAY78047;  
 XX 25-APR-2000 (first entry)  
 XX Human lactoferrin derived peptide SEQ ID NO:47.  
 DE Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200001730-A1.  
 PN 13-JAN-2000.  
 PD 06-JUL-1999; 99WO-SE01230.  
 PF 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX

PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 18; Page 73; 102pp; English.  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX SQ Sequence 12 AA;  
 Query Match 83.8%; Score 57; DB 21; Length 12;  
 Best Local Similarity 83.3%; Pred. No. 0.0039;  
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CFQWQRNMRKVR 12  
 Db |||||  
 1 CFQWQRNMRKVR 12  
 RESULT 12  
 AAY78037  
 ID AAY78037 standard; Peptide; 13 AA.  
 AC AAY78037;  
 XX 25-APR-2000 (first entry)  
 XX Human lactoferrin derived peptide SEQ ID NO:37.  
 DE Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200001730-A1.  
 PN 13-JAN-2000.  
 PD 06-JUL-1999; 99WO-SE01230.  
 PF 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX

BS Claim 12; Page 70; 102pp; English.

CC AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through

CC binding to specific lactoferrin receptors and are then transported

CC through the circulation. A medicinal product of the peptide or fragment

CC can be used for treating and/or prevention of infections (such as

CC urinary tract infections, colitis, and Candida infection on a mucosal

CC membrane), inflammations and/or tumours. The peptides can also be used

CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.

CC Even though native human lactoferrin have been shown to have desired

CC anti-inflammatory anti-infectious and anti-tumoural properties they

CC cannot be used clinically on a broad basis because of high production

CC costs. Therefore, provision of peptides based on lactoferrin would

CC enable them to be used for the same purposes as lactoferrin at lower

CC cost.

SQ Sequence 13 AA;

Query Match 83.8%; Score 57; DB 21; Length 13;

Best Local Similarity 83.3%; Pred. No. 0.0042;

Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFAWQAMRKVR 12

DB 2 CFQWQNRKVR 13

RESULT 13

AAAY78048

ID AAY78048 standard; Peptide; 13 AA.

AC AAY78048;

DT 25-APR-2000 (first entry)

DE Human lactoferrin derived peptide SEQ ID NO:48.

DE Human; lactoferrin; modification; infection; inflammation; tumour;

KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;

KW urinary tract infection; colitis; Candida infection; fungicidal;

KW bactericidal; preservative.

OS Homo sapiens.

OS Synthetic.

PN WO200001730-A1.

PD 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

FA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

PI WPI; 2000-147389/13.

DR New peptides used for treatment and prevention of infections,

PT inflammations and tumors and for use in infant formula food -

XX Claim 15; Page 74; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through

CC binding to specific lactoferrin receptors and are then transported

CC through the circulation. A medicinal product of the peptide or fragment

CC can be used for treating and/or prevention of infections (such as

CC urinary tract infections, colitis, and Candida infection on a mucosal

CC membrane), inflammations and/or tumours. The peptides can also be used

CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.

CC Even though native human lactoferrin have been shown to have desired

CC anti-inflammatory anti-infectious and anti-tumoural properties they

CC cannot be used clinically on a broad basis because of high production

CC costs. Therefore, provision of peptides based on lactoferrin would

CC enable them to be used for the same purposes as lactoferrin at lower

CC cost.

SQ Sequence 13 AA;

Query Match 83.8%; Score 57; DB 21; Length 13;

Best Local Similarity 83.3%; Pred. No. 0.0042;

Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFAWQAMRKVR 12

DB 2 CFQWQNRKVR 13

RESULT 14

AAAY78049

ID AAY78049 standard; Peptide; 13 AA.

AC AAY78049;

DT 25-APR-2000 (first entry)

DE Human lactoferrin derived peptide SEQ ID NO:49.

DE Human; lactoferrin; modification; infection; inflammation; tumour;

KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;

KW urinary tract infection; colitis; Candida infection; fungicidal;

KW bactericidal; preservative.

OS Homo sapiens.

OS Synthetic.

PN WO200001730-A1.

PD 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

FA Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

PI WPI; 2000-147389/13.

DR New peptides used for treatment and prevention of infections,

PT inflammations and tumors and for use in infant formula food -

XX Claim 18; Page 74; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through

CC binding to specific lactoferrin receptors and are then transported

CC through the circulation. A medicinal product of the peptide or fragment

CC can be used for treating and/or prevention of infections (such as

CC urinary tract infections, colitis, and Candida infection on a mucosal

CC membrane), inflammations and/or tumours. The peptides can also be used

CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.

CC Even though native human lactoferrin have been shown to have desired

CC anti-inflammatory anti-infectious and anti-tumoural properties they

CC cannot be used clinically on a broad basis because of high production

CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 13 AA;

Query Match 83.8%; Score 57; DB 21; Length 13;

Best Local Similarity 83.3%; Pred. No. 0.0042;

Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFAWQAMRKVR 12

Db 2 CFQWQNRKVR 13

# RESULT 15

AAAY78036

ID AAY78036 standard; Peptide; 14 AA.

XX AC

AAAY78036;

XX AC

25-APR-2000 (first entry)

XX DT

Human lactoferrin derived peptide SEQ ID NO:36.

XX DE

Human; lactoferrin; modification; infection; inflammation; tumour;  
 food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 urinary tract infection; colitis; Candida infection; fungicidal;  
 bactericidal; preservative.

XX KW

Homo sapiens.

OS OS

Synthetic.

XX PN

WO200001730-A1.

XX PD

13-JAN-2000.

XX PF

06-JUL-1999; 99WO-SE01230.

XX PR

06-JUL-1998; 98SE-0002441.

XX PR

17-JUL-1998; 98SE-0002562.

XX PR

29-DEC-1998; 98SE-0004614.

XX PA

(ASCI-) A+ SCI INVEST AB.

XX PI

Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX DR

WPI; 2000-147388/13.

XX PT

New peptides used for treatment and prevention of infections,

XX PT

inflammations and tumors and for use in infant formula food

XX PS

Claim 12; Page 69; 102pp; English.

XX CC

AAAY78001 to AAY78100 represent peptides having sequences based on human  
 lactoferrin. The peptides are taken up in the intestine through  
 binding to specific lactoferrin receptors and are then transported  
 through the circulation. A medicinal product of the peptide or fragment  
 can be used for treating and/or prevention of infections (such as  
 urinary tract infections, colitis, and candida infection on a mucosal  
 membrane), inflammations and/or tumours. The peptides can also be used  
 in food stuffs such as infant formula food. The peptides are also  
 fungicidal and bactericidal and may also be used as preservatives.  
 Even though native human lactoferrin have been shown to have desired  
 anti-inflammatory anti-infectious and anti-tumoural properties they  
 cannot be used clinically on a broad basis because of high production  
 costs. Therefore, provision of peptides based on lactoferrin would  
 enable them to be used for the same purposes as lactoferrin at lower  
 cost.

XX SQ Sequence 14 AA;

Query Match

83.8%; Score 57; DB 21; Length 14;

Best Local Similarity 83.3%; Pred. No. 0.0045;

Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFAWQAMRKVR 12

Db 3 CFQWQNRKVR 14

Search completed: February 21, 2003, 07:56:45

Job time : 29.093 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: February 21, 2003, 07:50:40 ; Search time 8.93023 Seconds  
(without alignments)  
39.537 Million cell updates/sec

Title: US-09-743-107b-95  
Perfect score: 68  
Sequence: 1 CFAWQRANKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	83.8	18	1	US-08-204-487-3
2	57	83.8	18	2	US-08-485-948-8
3	57	83.8	18	2	US-08-528-380-8
4	57	83.8	18	2	US-08-475-055-8
5	57	83.8	20	1	US-07-755-161A-3
6	57	83.8	20	1	US-07-891-174-3
7	57	83.8	20	1	US-08-204-487-1
8	57	83.8	20	1	US-08-256-771-24
9	57	83.8	20	1	US-08-256-771-25
10	57	83.8	20	1	US-08-381-984-24
11	57	83.8	20	1	US-08-381-984-25
12	57	83.8	22	4	US-09-508-734-4
13	57	83.8	24	4	US-09-508-734-6
14	57	83.8	25	1	US-07-755-161A-10
15	57	83.8	25	1	US-07-891-174-10
16	57	83.8	25	1	US-08-204-487-7
17	57	83.8	29	4	US-09-508-734-8
18	57	83.8	36	1	US-07-755-161A-8
19	57	83.8	36	1	US-07-891-174-8
20	57	83.8	36	1	US-08-256-771-30
21	57	83.8	36	1	US-08-381-984-29
22	57	83.8	47	2	US-08-464-182A-6
23	57	83.8	47	2	US-08-406-271-6
24	57	83.8	50	2	US-08-693-274A-7
25	57	83.8	52	4	US-09-017-043A-3
26	57	83.8	53	2	US-08-464-182A-5
27	57	83.8	53	2	US-08-406-271-5

28	57	83.8	54	2	US-08-464-182A-2	Sequence 2, Appli
29	57	83.8	54	2	US-08-406-271-2	Sequence 2, Appli
30	57	83.8	694	3	US-08-724-586-2	Sequence 2, Appli
31	57	83.8	694	4	US-09-421-632-2	Sequence 2, Appli
32	57	83.8	694	4	US-09-932-190-2	Sequence 2, Appli
33	57	83.8	705	2	US-08-655-640-2	Sequence 2, Appli
34	57	83.8	708	2	US-08-655-640-4	Sequence 4, Appli
35	57	83.8	711	1	US-08-154-019-4	Sequence 4, Appli
36	57	83.8	711	1	US-08-461-333-4	Sequence 4, Appli
37	57	83.8	711	3	US-08-464-167-4	Sequence 4, Appli
38	57	83.8	711	3	US-09-158-313-4	Sequence 4, Appli
39	57	83.8	711	4	US-08-476-798-4	Sequence 4, Appli
40	54	79.4	711	1	US-08-145-681-2	Sequence 2, Appli
41	54	79.4	711	1	US-08-250-308-2	Sequence 2, Appli
42	54	79.4	711	1	US-08-453-703-2	Sequence 2, Appli
43	54	79.4	711	2	US-08-456-106-2	Sequence 2, Appli
44	54	79.4	711	3	US-08-456-108-2	Sequence 2, Appli
45	54	79.4	711	4	US-09-265-577-2	Sequence 2, Appli

## ALIGNMENTS

RESULT 1  
US-08-204-487-3  
; Sequence 3, Application US/08204487  
; Patent No. 5565425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAOKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGEKI  
; APPLICANT: DOSAKO, SHUN'ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; TITLE OF INVENTION: INHIBITORS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESSEE: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FJN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..18  
; OTHER INFORMATION: /note= "HUMAN LACTOFERRIN PEPTIDE"



OTHER INFORMATION: (20-37) \*  
US-08-204-487-3

Query Match 83.8%; Score 57; DB 1; Length 18;  
Best Local Similarity 83.3%; Pred. No. 0.0016;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12  
| | | | |  
DB 1 CFQWQRMKVR 12

## RESULT 2

US-08-485-948-8  
; Sequence 8, Application US/08485948  
; Patent No. 5855882  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VIASSARA  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,948  
; FILING DATE:  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/488,217  
; FILING DATE: JUNE 7, 1995  
; APPLICATION NUMBER: 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LF-Cl, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
US-08-485-948-8

Query Match 83.8%; Score 57; DB 2; Length 18;  
Best Local Similarity 83.3%; Pred. No. 0.0016;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12  
| | | | |  
DB 1 CFQWQRMKVR 12

## RESULT 3

US-08-628-380-8  
; Sequence 8, Application US/08628380  
; Patent No. 5891341  
; GENERAL INFORMATION:  
; APPLICANT: LI, YONG MING  
; APPLICANT: VIASSARA, HELEN  
; APPLICANT: CERAMI, ANTHONY  
; TITLE OF INVENTION: AGE-MEDIATED INHIBITION OF ANTIBACTERIAL PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/628,380  
; FILING DATE: April 4, 1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LF-Cl, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal  
US-08-628-380-8

Query Match 83.8%; Score 57; DB 2; Length 18;  
Best Local Similarity 83.3%; Pred. No. 0.0016;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12  
| | | | |  
DB 1 CFQWQRMKVR 12

## RESULT 4

US-08-475-055-8  
; Sequence 8, Application US/08475055  
; Patent No. 5962245  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VIASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,055
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/485,948
FILING DATE:
APPLICATION NUMBER: 08/488,217
FILING DATE: JUNE 7, 1995
APPLICATION NUMBER: 08/418,642
FILING DATE: APRIL 7, 1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 947-1-008A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
DESCRIPTION: LF-Cl, 8-25
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
US-08-475-055-8

Query Match      83.8%; Score 57; DB 2; Length 18;
Best Local Similarity 83.3%; Pred. No. 0.0016;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQWRMRKVR 12
Db 1 CFQWRMRKVR 12

RESULT 5
US-07-755-161A-3
Sequence 3, Application US/07755161A
Patent No. 5304633
GENERAL INFORMATION:
APPLICANT: Mamoru TOMITA et al.
TITLE OF INVENTION: Antimicrobial Peptides and an
TITLE OF INVENTION: Antimicrobial Agent
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
COMPUTER: IBM Compatible

```

```

OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/755,161A
FILING DATE: 19910905
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX: 202-371-8856
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY: modified site
LOCATION: 2
IDENTIFICATION METHOD:
OTHER INFORMATION: /note= "thiol group of
OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with
OTHER INFORMATION: thiol group of Cys residue at location 19"
FEATURE:
NAME/KEY: modified site
LOCATION: 19
IDENTIFICATION METHOD:
OTHER INFORMATION: /note= "thiol group of
OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with
OTHER INFORMATION: thiol group of Cys residue at location 2"
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-07-755-161A-3

```

Query Match

83.8%; Score 57; DB 1; Length 20;

```
Best Local Similarity 83.3%; Pred. No. 0.0017;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFAMQRMKVR 12
Db 2 CFQWRNMRKVR 13

RESULT 6
US-07-891-174-3
; Sequence 3, Application US/07891174
; Patent No. 5317084
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; TITLE OF INVENTION: Antimicrobial Agent
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/891,174
; FILING DATE: 29-MAY-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/755,161
; FILING DATE: 05-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site

; LOCATION: 2
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 19"
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 19
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 2"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-07-891-174-3

Query Match 83.8%; Score 57; DB 1; Length 20;
Best Local Similarity 83.3%; Pred. No. 0.0017;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFAMQRMKVR 12
Db 2 CFQWRNMRKVR 13

RESULT 7
US-08-204-487-1
; Sequence 1, Application US/08204487
; Patent No. 5565425
; GENERAL INFORMATION:
; APPLICANT: YAMAMOTO, NAKI
; APPLICANT: NAKASHIMA, HIDEKI
; APPLICANT: MOSUCHI, WATARU
; APPLICANT: TANAKA, SHIGEKI
; APPLICANT: DOSAKO, SHUN'ICHI
; APPLICANT: KAWASAKI, YOSHIHIRO
; APPLICANT: UCHIDA, TOSHIKI
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION
; TITLE OF INVENTION: INHIBITORS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
; ADDRESSEE: THIBEAULT
; STREET: 53 STATE STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/204,487
; FILING DATE: 02-MAR-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, PAULA A.
; REGISTRATION NUMBER: 32,503
; REFERENCE/DOCKET NUMBER: FJN-019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
```

TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..20  
OTHER INFORMATION: /note= "ANTIBACTERIAL PEPTIDE  
DERIVED FROM HUMAN LACTOPERRIN"  
US-08-204-487-1

Query Match 83.8%; Score 57; DB 1; Length 20;  
Best Local Similarity 83.3%; Pred. No. 0.0017; 2; Indels 0;  
Matches 10; Conservative 0; Mismatches 2; Gaps 0;

Qy 1 CFAWQAMRKVR 12  
||| ||| ||| |||  
Db 2 CFWQRMNRKVR 13

RESULT 8  
US-08-256-771-24  
; Sequence 24, Application US/08256771  
; Patent No. 5656591  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
; PRODUCT'S THEREWITH  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,771  
FILING DATE: July 22, 1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:

INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "Cys residues are linked by  
disulfide bond"  
OTHER INFORMATION:

US-08-256-771-24

Query Match 83.8%; Score 57; DB 1; Length 20;  
Best Local Similarity 83.3%; Pred. No. 0.0017; 2; Indels 0;  
Matches 10; Conservative 0; Mismatches 2; Gaps 0;

Qy 1 CFAWQAMRKVR 12  
||| ||| ||| |||  
Db 2 CFWQRMNRKVR 13

RESULT 9  
US-08-256-771-25  
; Sequence 25, Application US/08256771  
; Patent No. 5656591  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING  
; PRODUCT'S THEREWITH  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/256,771  
FILING DATE: July 22, 1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:

INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "Cys residues are protected to  
prevent disulfide bond"  
OTHER INFORMATION:  
US-08-256-771-25

Query Match 83.8%; Score 57; DB 1; Length 20;  
Best Local Similarity 83.3%; Pred. No. 0.0017;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFAWQAMRKVR 12  
||| ||| ||| |||  
Db 2 CFWQRMNRKVR 13

RESULT 10  
US-08-381-984-24

Sequence 24, Application US/08381984  
Patent No. 5804555  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIOXIDANT  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/381,984  
FILING DATE: April 11, 1995  
CLASSIFICATION: 252  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "cysteine residues at positions 2  
OTHER INFORMATION: and 19 are bonded by disulfide linkage"  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "the specified peptide as well as  
OTHER INFORMATION: peptides including the specified peptide as a fragment thereof  
US-08-381-984-24  
Query Match 83.8%; Score 57; DB 1; Length 20;  
Best Local Similarity 83.3%; Pred. No. 0.0017;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CFAWQRAVRKVR 12  
Db 2 CFQWQNRKVR 13  
RESULT 11  
US-08-381-984-25  
Sequence 25, Application US/08381984  
Patent No. 5804555  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIOXIDANT  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Wenderoth, Lind & Ponack

STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/381,984  
FILING DATE: April 11, 1995  
CLASSIFICATION: 252  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "the specified peptide as well as  
OTHER INFORMATION: peptides including the specified peptide as a fragment thereof  
US-08-381-984-25  
Query Match 83.8%; Score 57; DB 1; Length 20;  
Best Local Similarity 83.3%; Pred. No. 0.0017;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CFAWQRAVRKVR 12  
Db 2 CFQWQNRKVR 13  
RESULT 12  
US-09-508-734-4  
Sequence 4, Application US/09508734  
Patent No. 6423509  
GENERAL INFORMATION:  
APPLICANT: Samyang Genex Corporation  
TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast an  
TITLE OF INVENTION: useful microorganism thereof  
FILE REFERENCE: PA/SYG/00139  
CURRENT APPLICATION NUMBER: US/09/508,734  
CURRENT FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: PCT/KR99/00373  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: KR1998-29351  
PRIOR FILING DATE: 1998-07-13  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Kopatentin 1.71  
SEQ ID NO 4

```
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-508-734-4
  Query Match      83.8%; Score 57; DB 4; Length 22;
  Best Local Similarity 83.3%; Pred. No. 0.0019; 2; Indels 0; Gaps 0;
  Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQWQRMNRKVR 12
  |||||
Db 2 CFQWQRMNRKVR 13

RESULT 13
US-09-508-734-6
; Sequence 6, Application US/09508734
; Patent No. 6423509
; GENERAL INFORMATION:
; APPLICANT: Sanyang Genex Corporation
; TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and
; FILE REFERENCE: PA/SYG/00139
; CURRENT APPLICATION NUMBER: US/09/508,734
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/KR99/00373
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: KR1998-29351
; PRIOR FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Kopatentin 1.71
; SEQ ID NO 6
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-508-734-6

  Query Match      83.8%; Score 57; DB 4; Length 24;
  Best Local Similarity 83.3%; Pred. No. 0.002; 2; Indels 0; Gaps 0;
  Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQWQRMNRKVR 12
  |||||
Db 3 CFQWQRMNRKVR 14

RESULT 14
US-07-755-161A-10
; Sequence 10, Application US/07755161A
; Patent No. 5304633
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; TITLE OF INVENTION: Antimicrobial Agent
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Displaywrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/755,161A
; FILING DATE: 19910905
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
```

```
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURES:
; NAME/KEY: modified site
; LOCATION: 4
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 21"
; FEATURES:
; NAME/KEY: modified site
; LOCATION: 21
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 4"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-07-755-161A-10

  Query Match      83.8%; Score 57; DB 1; Length 25;
  Best Local Similarity 83.3%; Pred. No. 0.0021;
  Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFQWQRMNRKVR 12
  |||||
Db 4 CFQWQRMNRKVR 15
```

## RESULT 15

US-07-891-174-10  
Sequence 10, Application US/07891174  
Patent No. 5317084  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: Antimicrobial Peptides and an  
TITLE OF INVENTION: Antimicrobial Agent  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/891,174  
FILING DATE: 29-MAY-1992  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/755,161  
FILING DATE: 05-SEP-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX: 202-371-8856  
TELEX:  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 4  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 21"  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 21

## IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 4"  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-891-174-10  
Query Match 83.8%; Score 57; DB 1; Length 25;  
Best Local Similarity 83.3%; Pred. No. 0.0021;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CFAMQAMRKVR 12  
Db 4 CFQMQRNMRKVR 15  
Search completed: February 21, 2003, 08:04:27  
Job time : 8.93023 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 08:08:15 ; Search time 6.88372 Seconds  
(without alignments)  
54.162 Million cell updates/sec

Title: US-09-743-107B-95  
Perfect score: 68  
Sequence: 1 CFAPORAMEKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 156504 seqs, 31069816 residues

Total number of hits satisfying chosen parameters: 156504

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:\*
- 1: /cgn2\_6/ptodata/2/pubpaa/US08 NEW PUB.pcp.\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/PCT NEW PUB.pcp.\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US02 NEW PUB.pcp.\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06 PUBCOMB.pcp.\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07 NEW PUB.pcp.\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/US07 PUBCOMB.pcp.\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS PUBCOMB.pcp.\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US08 PUBCOMB.pcp.\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09 NEW PUB.pcp.\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US03 PUBCOMB.pcp.\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US10 NEW PUB.pcp.\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US10 PUBCOMB.pcp.\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US60 NEW PUB.pcp.\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US60 PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	83.8	15	9	US-09-798-869-2
2	57	83.8	25	9	US-09-798-869-20
3	57	83.8	694	9	US-10-023-096-2
4	49	72.1	15	9	US-09-798-869-6
5	47	69.1	15	9	US-09-798-869-3
6	47	69.1	25	9	US-09-798-869-23
7	42	61.8	15	9	US-09-798-869-4
8	42	61.8	25	9	US-09-798-869-22
9	39	57.4	15	9	US-09-798-869-7
10	39	57.4	15	9	US-09-798-869-8
11	39	57.4	15	9	US-09-798-869-29
12	39	57.4	15	9	US-09-798-869-30
13	38	55.9	138	9	US-10-013-379-14
14	27	54.4	2273	10	US-09-935-542-12
15	37	54.4	2310	10	US-09-935-542-10
16	36	52.9	210	9	US-09-738-626-4884
17	36	52.9	301	9	US-10-080-960-11
18	36	52.9	302	10	US-09-948-078-2
19	35	51.5	489	9	US-09-888-320-2

20	34	50.0	46	10	US-09-864-761-35744	Sequence 35744, A
21	34	50.0	69	10	US-09-804-969-17	Sequence 17, Appl
22	34	50.0	209	10	US-09-904-536-8	Sequence 8, Appl
23	34	50.0	209	10	US-09-904-536-9	Sequence 9, Appl
24	34	50.0	209	10	US-09-904-536-11	Sequence 11, Appl
25	34	50.0	209	10	US-09-904-536-12	Sequence 12, Appl
26	34	50.0	209	10	US-09-904-536-13	Sequence 13, Appl
27	34	50.0	209	10	US-09-904-536-14	Sequence 14, Appl
28	34	50.0	209	10	US-09-904-536-15	Sequence 15, Appl
29	34	50.0	209	10	US-09-904-536-16	Sequence 16, Appl
30	34	50.0	209	10	US-09-904-536-17	Sequence 17, Appl
31	34	50.0	209	10	US-09-904-536-18	Sequence 18, Appl
32	34	50.0	212	10	US-09-904-536-10	Sequence 10, Appl
33	34	50.0	235	9	US-10-095-449-6	Sequence 6, Appl
34	34	50.0	235	10	US-09-448-378-1	Sequence 1, Appl
35	34	50.0	235	10	US-09-983-806-6	Sequence 6, Appl
36	34	50.0	235	10	US-09-904-536-1	Sequence 1, Appl
37	34	50.0	365	9	US-10-109-533A-2	Sequence 2, Appl
38	34	50.0	372	9	US-09-764-864-1044	Sequence 1044, Ap
39	34	50.0	1192	9	US-10-170-102-2	Sequence 2, Appl
40	33	48.5	67	9	US-09-796-632-1487	Sequence 1487, Ap
41	33	48.5	81	10	US-09-864-761-41887	Sequence 41887, A
42	33	48.5	171	10	US-09-764-864-1119	Sequence 1119, Ap
43	33	48.5	171	10	US-09-764-864-1536	Sequence 1536, Ap
44	33	48.5	579	10	US-09-824-588-2	Sequence 2, Appl
45	33	48.5	620	10	US-09-764-864-1116	Sequence 1116, Ap

ALIGNMENTS

RESULT 1  
US-09-798-869-2  
; Sequence 2, Application US/09798869  
; Publication No. US2003002821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LAES VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798.869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-09-798-869-2

Query Match 83.8%; Score 57; DB 9; Length 15;  
Best Local Similarity 83.3%; Pred. No. 0.00055;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFAPORAMEKVR 12  
DB 3 CFQWQNRKVR 14

RESULT 2  
US-09-798-869-20  
; Sequence 20, Application US/09798869  
; Publication No. US2003002821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON



; APPLICANT: LARS VORLAND  
 ; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
 ; FILE REFERENCE: A34049-PCT-USA-A  
 ; CURRENT APPLICATION NUMBER: US/09/798,869  
 ; CURRENT FILING DATE: 2001-02-27  
 ; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
 ; PRIOR FILING DATE: 1999-08-31  
 ; PRIOR APPLICATION NUMBER: GB9818938.4  
 ; PRIOR FILING DATE: 1998-08-28  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 20  
 ; LENGTH: 25  
 ; TYPE: PRT  
 ; ORGANISM: Homo Sapiens  
 US-09-798-869-20

Query Match 83.8%; Score 57; DB 9; Length 25;  
 Best Local Similarity 83.3%; Pred. No. 0.0093;  
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12  
 |||||  
 Db 3 CFQWQRMKVR 14

RESULT 3  
 US-10-023-096-2  
 ; Sequence 2, Application US/10023096  
 ; Patent No. US20020160941A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kruzel, Marian L.  
 ; APPLICANT: Kurecki, Tomasz  
 ; APPLICANT: Gollnick, Paul D.  
 ; APPLICANT: Doyle, Darrell J.  
 ; TITLE OF INVENTION: Cloning, Expression, and Uses of Human  
 ; TITLE OF INVENTION: Lactoferrin  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Jacobson, Price, Holman & Stern  
 ; STREET: 400 Seventh St. N.W.  
 ; CITY: Washington D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20004

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/10/023,096  
 ; FILING DATE:

; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/724,586  
 ; FILING DATE: 30-SEPT-1996  
 ; APPLICATION NUMBER: US 08/238,445  
 ; FILING DATE: 05-MAY-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Player, William E.  
 ; REGISTRATION NUMBER: 31,409  
 ; REFERENCE/DOCKET NUMBER: 10505/P58185C  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202) 638-6666  
 ; TELEFAX: (202) 393-9350  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 694 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 US-10-023-096-2

Query Match 83.8%; Score 57; DB 9; Length 694;  
 Best Local Similarity 83.3%; Pred. No. 0.027;  
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12  
 |||||  
 Db 22 CFQWQRMKVR 33

RESULT 4  
 US-09-798-869-6  
 ; Sequence 6, Application US/09798869  
 ; Publication No. US20030022821A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOHN SIGURD SVENDSEN  
 ; APPLICANT: (VSTEIN REKDAL  
 ; APPLICANT: BALDUR SVEINBJ (RNSSON  
 ; APPLICANT: LARS VORLAND  
 ; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
 ; FILE REFERENCE: A34049-PCT-USA-A  
 ; CURRENT APPLICATION NUMBER: US/09/798,869  
 ; CURRENT FILING DATE: 2001-02-27  
 ; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
 ; PRIOR FILING DATE: 1999-08-31  
 ; PRIOR APPLICATION NUMBER: GB9818938.4  
 ; PRIOR FILING DATE: 1998-08-28  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 6  
 ; LENGTH: 15  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of Homo sapiens  
 ; OTHER INFORMATION: sequence)  
 US-09-798-869-6

Query Match 72.1%; Score 49; DB 9; Length 15;  
 Best Local Similarity 75.0%; Pred. No. 0.013;  
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12  
 |||||  
 Db 3 CFQWQRMKVR 14

RESULT 5  
 US-09-798-869-3  
 ; Sequence 3, Application US/09798869  
 ; Publication No. US20030022821A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOHN SIGURD SVENDSEN  
 ; APPLICANT: (VSTEIN REKDAL  
 ; APPLICANT: BALDUR SVEINBJ (RNSSON  
 ; APPLICANT: LARS VORLAND  
 ; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
 ; FILE REFERENCE: A34049-PCT-USA-A  
 ; CURRENT APPLICATION NUMBER: US/09/798,869  
 ; CURRENT FILING DATE: 2001-02-27  
 ; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
 ; PRIOR FILING DATE: 1999-08-31  
 ; PRIOR APPLICATION NUMBER: GB9818938.4  
 ; PRIOR FILING DATE: 1998-08-28  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3  
 ; LENGTH: 15  
 ; TYPE: PRT  
 ; ORGANISM: CAPRINE  
 US-09-798-869-3

Query Match 69.1%; Score 47; DB 9; Length 15;  
 Best Local Similarity 63.6%; Pred. No. 0.029;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFAWQAMRKV 11  
|:|||||  
Db 3 CYQWRMRKL 13

## RESULT 6

US-09-798-869-23  
; Sequence 23, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: CAPRINE  
US-09-798-869-23

Query Match 69.1%; Score 47; DB 9; Length 25;  
Best Local Similarity 63.6%; Pred. No. 0.049;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFAWQAMRKV 11  
|:|||||  
Db 3 CYQWRMRKL 13

## RESULT 7

US-09-798-869-4  
; Sequence 4, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-798-869-4

Query Match 61.8%; Score 42; DB 9; Length 15;  
Best Local Similarity 63.6%; Pred. No. 0.21;  
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFAWQAMRKV 11  
|:|||||  
Db 3 CLRQWEMRKV 13

## RESULT 8

US-09-798-869-22  
; Sequence 22, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 22  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: MURINE  
US-09-798-869-22

Query Match 61.8%; Score 42; DB 9; Length 25;  
Best Local Similarity 63.6%; Pred. No. 0.35;  
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFAWQAMRKV 11  
|:|||||  
Db 3 CLRQWEMRKV 13

## RESULT 9

US-09-798-869-7  
; Sequence 7, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic peptide (modified form of homo sapiens  
; OTHER INFORMATION: sequence)  
US-09-798-869-7

Query Match 57.4%; Score 39; DB 9; Length 15;  
Best Local Similarity 54.5%; Pred. No. 0.69;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFAWQAMRKV 11  
|:|||||  
Db 3 CYQWRMRKL 13

```
RESULT 10
US-09-798-869-8
; Sequence 8, Application US/09798869
; Publication No. US20030022821A1
; GENERAL INFORMATION:
; APPLICANT: JOHN SIGURD SVENDSEN
; APPLICANT: (YSTEIN REKDAL
; APPLICANT: BALDUR SVEINBJ (RNSSON
; APPLICANT: LARS VORLAND
; TITLE OF INVENTION: BIOACTIVE PEPTIDES
; FILE REFERENCE: A34049-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/798,869
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: PCT/GB99/02851
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: GB9818938.4
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 15
; TYPE: PRT
; ORGANISM: BOVINE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of caprine
; OTHER INFORMATION: sequence)
US-09-798-869-8

Query Match          57.4%; Score 39; DB 9; Length 15;
Best Local Similarity 63.6%; Pred. No. 0.69;
Matches 7; Conservative 0; Mismatches 4; Indels 4; Gaps 0;

QY 1 CFAPQAMRKV 11
DB 3 CLRQWQMRKVL 13

RESULT 11
US-09-798-869-29
; Sequence 29, Application US/09798869
; Publication No. US20030022821A1
; GENERAL INFORMATION:
; APPLICANT: JOHN SIGURD SVENDSEN
; APPLICANT: (YSTEIN REKDAL
; APPLICANT: BALDUR SVEINBJ (RNSSON
; APPLICANT: LARS VORLAND
; TITLE OF INVENTION: BIOACTIVE PEPTIDES
; FILE REFERENCE: A34049-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/798,869
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: PCT/GB99/02851
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: GB9818938.4
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 15
; TYPE: PRT
; ORGANISM: BOVINE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of caprine
; OTHER INFORMATION: sequence)
US-09-798-869-29

Query Match          57.4%; Score 39; DB 9; Length 15;
Best Local Similarity 54.5%; Pred. No. 0.69;
Matches 6; Conservative 2; Mismatches 3; Indels 3; Gaps 0;

QY 1 CFAPQAMRKV 11
DB 3 CLRQWQMRKVL 13

RESULT 12
US-09-798-869-30
; Sequence 30, Application US/09798869
; Publication No. US20030022821A1
; GENERAL INFORMATION:
; APPLICANT: JOHN SIGURD SVENDSEN
; APPLICANT: (YSTEIN REKDAL
; APPLICANT: BALDUR SVEINBJ (RNSSON
; APPLICANT: LARS VORLAND
; TITLE OF INVENTION: BIOACTIVE PEPTIDES
; FILE REFERENCE: A34049-PCT-USA-A
; CURRENT APPLICATION NUMBER: US/09/798,869
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: PCT/GB99/02851
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: GB9818938.4
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 15
; TYPE: PRT
; ORGANISM: BOVINE
; FEATURE:
; OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of caprine
; OTHER INFORMATION: sequence)
US-09-798-869-30

Query Match          57.4%; Score 39; DB 9; Length 15;
Best Local Similarity 54.5%; Pred. No. 0.69;
Matches 6; Conservative 2; Mismatches 3; Indels 3; Gaps 0;

QY 1 CFAPQAMRKV 11
DB 3 CLRQWQMRKVL 13

RESULT 13
US-10-013-379-14
; Sequence 14, Application US/10013379
; Publication No. US20020188108A1
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of California
; APPLICANT: Cate, Jamie H.
; APPLICANT: No. US20020188108A1, Harry F.
; APPLICANT: Yusupov, Marat M.
; APPLICANT: Yusupova, Guinara ZH
; APPLICANT: Baucum, Albion
; APPLICANT: Lancaster, Laura
; APPLICANT: Dallas, Anne
; TITLE OF INVENTION: X-RAY CRYSTAL STRUCTURES OF FUNCTIONAL
; COMPLEXES OF THE BACTERIAL RIBOSOME CONTAINING TRANSFER RNA
; TITLE OF INVENTION: AND MODEL MESSENGER RNAs
; FILE REFERENCE: 19629-7010
; CURRENT APPLICATION NUMBER: US/10/013,379
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: US 60/254,603
; PRIOR FILING DATE: 2000-12-09
; PRIOR APPLICATION NUMBER: US 60/278,013
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/294,394
; PRIOR FILING DATE: 2001-05-30
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Thermus thermophilus
; FEATURE:
; OTHER INFORMATION: 50S ribosomal protein L16
; OTHER INFORMATION: 19iyp
US-10-013-379-14

Query Match          55.9%; Score 38; DB 9; Length 138;
Best Local Similarity 70.0%; Pred. No. 9.9;
Matches 7; Conservative 1; Mismatches 2; Indels 2; Gaps 0;

QY 2 FAPQAMRKV 11
```

db 115 FAWRAYMKV 124  
|||||

## RESULT 14

JS-09-995-542-12  
; Sequence 12, Application US/09995542  
; Patent No. US20020127647A1  
; GENERAL INFORMATION:  
; APPLICANT: Shutter, John  
; APPLICANT: Ullias, Learni  
; TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and  
; TITLE OF INVENTION: Uses Thereof  
; FILE REFERENCE: 00-658-A  
; CURRENT APPLICATION NUMBER: US/09/995,542  
; CURRENT FILING DATE: 2001-11-28  
; PRIOR APPLICATION NUMBER: 60/253,520  
; PRIOR FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 2273  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
JS-09-995-542-12

Query Match 54.4%; Score 37; DB 10; Length 2273;  
Best Local Similarity 75.0%; Pred. No. 2.5e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

2Y 1 CFAWQRAM 8  
|||||  
db 779 CFAWQDRM 786

## RESULT 15

JS-09-995-542-10  
; Sequence 10, Application US/09995542  
; Patent No. US20020127647A1  
; GENERAL INFORMATION:  
; APPLICANT: Shutter, John  
; APPLICANT: Ullias, Learni  
; TITLE OF INVENTION: ATP-Binding Cassette Transporter-Like Molecules and  
; TITLE OF INVENTION: Uses Thereof  
; FILE REFERENCE: 00-658-A  
; CURRENT APPLICATION NUMBER: US/09/995,542  
; CURRENT FILING DATE: 2001-11-28  
; PRIOR APPLICATION NUMBER: 60/253,520  
; PRIOR FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 10  
; LENGTH: 2310  
; TYPE: PRT  
; ORGANISM: Mus musculus  
JS-09-995-542-10

Query Match 54.4%; Score 37; DB 10; Length 2310;  
Best Local Similarity 75.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

2Y 1 CFAWQRAM 8  
|||||  
db 779 CFAWQDRM 786

Search completed: February 21, 2003, 08:11:58  
Job time : 7.88372 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:48:01 ; Search time 10.6047 Seconds  
(without alignments)  
108.784 Million cell updates/sec

Title: US-09-743-107b-95

Perfect score: 68

Sequence: 1 CFAWQSAWKVR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR73:\*\*

1: PIR1:\*\*

2: PIR2:\*\*

3: PIR3:\*\*

4: PIR4:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	57	93.8	711	1	TFHUL	lactotransferrin p
2	47	69.1	708	2	JC2323	lactoferrin - goat
3	44	64.7	33	2	SS2107	lactoferrin - sheep
4	43	63.2	500	2	G71633	ADP,ATP carrier pr
5	42	61.8	121	2	AH3147	hypothetical prote
6	42	61.8	707	1	A28438	lactoferrin precu
7	40	58.8	114	2	D33876	carcinoembryonic a
8	40	58.8	306	1	A39654	cell cycle arrest
9	39	57.4	511	2	AB0858	hypothetical prote
10	39	57.4	1135	2	T14803	phytochrome C - so
11	38	55.9	289	2	G86403	33.3K hypothetical
12	38	55.9	161	2	B82395	methy-accepting c
13	37	54.4	161	2	AC3242	NTP pyrophosphohy
14	37	54.4	202	2	C71306	hypothetical prote
15	37	54.4	206	2	H97451	pyridoxamine 5'-ph
16	37	54.4	206	2	AB2670	pyridoxamine 5'-ph
17	37	54.4	208	2	AG3441	Probable pyridoxam
18	37	54.4	275	1	JCL113	interleukin-2 rece
19	37	54.4	275	1	S07442	interleukin-2 rece
20	37	54.4	294	2	T00104	probable drpp-4-de
21	37	54.4	303	1	WZBEM6	gene 20 protein -
22	37	54.4	303	2	C33374	hypothetical prote
23	37	54.4	337	2	B97374	Dainococcus radiod
24	37	54.4	337	2	AH2591	membrane lipoprote
25	37	54.4	494	2	F83199	probable carbonydr
26	37	54.4	499	2	B97842	ADP,ATP carrier pr
27	37	54.4	502	2	T01179	hypothetical prote
28	37	54.4	584	2	C84325	hypothetical prote
29	37	54.4	601	2	JEO238	stress protein p66

30 37 54.4 1636 2 B82736 hemolysin-type cal  
31 37 54.4 2700 2 D88450 protein F2IH1.2 [  
32 37 54.4 4568 2 T08030 dynein beta heavy  
33 36.5 53.7 278 2 AD0631 probable membrane  
34 36 52.9 396 2 E90238 threonine synthase  
35 36 52.9 456 2 C86624 hypothetical prote  
36 36 52.9 456 2 H72000 hypothetical prote  
37 36 52.9 541 2 C81357 probable secreted  
38 36 52.9 807 1 WWAD15 late 100K protein  
39 36 52.9 956 2 B71250 valine-tRNA ligase  
40 36 52.9 1081 2 T31329 receptor tyrosine  
41 36 52.9 1165 2 D59433 C. elegans protein  
42 36 52.9 1746 2 D83181 hypothetical prote  
43 35 51.5 228 2 AS0438 conserved hypotet  
44 35 51.5 243 2 S25308 probable phycocyan  
45 35 51.5 275 2 T22597 hypothetical prote

#### RESULT 1

TFHUL

lactotransferrin precursor [validated] - human

N:Alternate names: lactoferrin

C:Species: Homo sapiens (man)

C>Date: 31-Mar-1992 #sequence revision 21-Nov-1997 #text change 08-Dec-2000

C:Accession: G01394; S11228; A45401; S10324; S15853; S20841; S07160; A61189; A31000; S7  
R;Cho, Y.

submitted to the EMBL Data Library, March 1994

A:Reference number: G06820

A:Accession: G01394

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-711 <CH>

A:Cross-references: EMBL:U07643; NID:G467236; PIDN:AAB60324.1; PID:G467237

R:Rey, M.W.; Woloshuk, S.L.; deBoer, H.A.; Pieper, F.R.

Nucleic Acids Res. 18, 5288, 1990

A:Title: Complete nucleotide sequence of human mammary gland lactoferrin.

A:Reference number: S11228; MUID:90384839; PMID:2402455

A:Accession: S11228

A:Molecule type: mRNA

A:Residues: 1-148, 'T', 150-422, 'C', 424-711 <REY>

A:Cross-references: EMBL:X53961; NID:G34415; PIDN:CAA37914.1; PID:G34416

R:Teng, C.T.; Liu, Y.; Yang, N.; Walmer, D.; Panella, T.

Mol. Endocrinol. 6, 1969-1981, 1992

A:Title: Differential molecular mechanism of the estrogen action that regulates lactofe

A:Reference number: A45401; MUID:93125571; PMID:1480183

A:Accession: A45401

A:Molecule type: DNA

A:Residues: 1-15 <TEN>

A:Cross-references: GB:S52659; NID:G263311; PIDN:AAB24877.1; PID:G263312

A:Experimental source: placenta

A:Notes: sequence extracted from NCBI backbone (NCBI:P:122202)

R:Powell, M.J.; Ogden, J.E.

Nucleic Acids Res. 18, 4013, 1990

A:Title: Nucleotide sequence of human lactoferrin cDNA.

A:Reference number: S10324; MUID:90326549; PMID:2374734

A:Accession: S10324

A:Molecule type: mRNA

A:Residues: 3-711 <POW>

A:Cross-references: EMBL:X52941; NID:G34411; PIDN:CAA37116.1; PID:G34412

R:Stowell, K.M.; Rado, T.A.; Funk, W.D.; Tweedie, J.W.

Biochem. J. 276, 349-355, 1991

A:Title: Expression of cloned human lactoferrin in baby-hamster kidney cells.

A:Reference number: S15853; MUID:91264786; PMID:2049066

A:Accession: S15853

A>Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 20-31 <STI>

A:Accession: S20841

A:Molecule type: protein

A:Residues: 20-28, 'X', 30-31 <ST2>

```

R;Rado, T.A.; Wei, X.; Benz Jr., E.J.
Blood 70, 989-993, 1987
A;Title: Isolation of lactoferrin cDNA from a human myeloid library and expression of mRNA
A;Reference number: S07160; MUID:88001031; PMID:3477300
A;Accession: S07160
A;Molecule type: mRNA
A;Residues: 436-487, 'A', 489-711 <RAD>
A;Cross-references: EMBL:M18642; NID:9186815; PIDN:AAA8665.1; PID:9386955
R;Panella, T.J.; Liu, Y.; Huang, A.T.; Teng, C.T.
Cancer Res. 51, 3037-3043, 1991
A;Title: Polymorphism and altered methylation of the lactoferrin gene in normal leukocytes
A;Reference number: A61169; MUID:91235214; PMID:1674448
A;Accession: A61169
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 3-701, 'SWKPNV' <PAN>
A;Experimental source: normal breast tissue
R;Metz-Boutigue, M.H.; Jolles, J.; Mazurier, J.; Schoentgen, F.; Legrand, D.; Spik, G.;
Eur. J. Biochem. 145, 659-666, 1984
A;Title: Human lactotransferrin: amino acid sequence and structural comparisons with other
A;Reference number: A31000; MUID:85076657; PMID:6510420
A;Accession: A31000
A;Molecule type: protein
A;Residues: 20-140, 142-169, 171-203, 'L', 205, 'K', 207-208, 'K', 210-385, 'Q', 387-391, 'W', 393-4
A;Note: this is the final paper in a series
R;Houen, G.; Hoegdall, E.V.; Barkholt, V.; Nørskov, L.
Eur. J. Biochem. 241, 303-308, 1996
A;Title: Lactoferrin: similarity to diamine oxidase and purification by aminoethyl affinity
A;Reference number: S74119; MUID:97054624; PMID:8898921
A;Accession: S74119
A;Molecule type: protein
A;Residues: 'G', 23-24, 'R', 26-27, 'XX', 30-32 <HOU>
A;Experimental source: neutrophil granulocytes
C;Genetics:
A;Gene: GDB:LTF
A;Cross-references: GDB:119368; OMIM:150210
A;Map position: 3q21-3q23
C;Superfamily: transferrin; transferrin repeat homology
C;Keywords: duplication; glycoprotein; iron binding; milk
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-711/Product: lactotransferrin #status experimental <MAT>
F;21-356/Domain: transferrin repeat homology <TRH>
F;360-699/Domain: transferrin repeat homology <TRH2>
F;7360-699/Domain: transferrin repeat homology <TRH2>
F;157-498/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;368-400, 378-391, 425-706, 447-669, 479-554, 513-527, 524-537, 647-652/Disulfide bonds: #stab
Query Match 83.8%; Score 57; DB 1; Length 711;
Best Local Similarity 83.3%; Pred. No. 0.025; Mismatches 2; Indels 0; Gaps 0;
Matches 10; Conservative 0;
Qy 1 CFAWQAMRKVR 12
Db 39 CFQWQNRKVR 50

Query Match 69.1%; Score 47; DB 2; Length 708;
Best Local Similarity 63.6%; Pred. No. 1.5; Mismatches 2; Indels 0; Gaps 0;
Matches 7; Conservative 2;
Qy 1 CFAWQAMRKV 11
Db 38 CYQWQRMKRL 48

RESULT 3
S52107
lactoferrin - sheep (fragment)
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 14-Jul-1995 #sequence_revision 05-Dec-1997 #text_change 05-Dec-1997
C;Accession: S52107
R;Qian, Z.Y.; Jolles, P.; Migliore-Samouri, D.; Fiat, A.M.
Biochim. Biophys. Acta 1243, 25-32, 1995
A;Title: Isolation and characterization of sheep lactoferrin, an inhibitor of platelet
A;Reference number: S52107; MUID:95127729; PMID:7827104
A;Accession: S52107
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-33 <QIA>
C;Superfamily: transferrin; transferrin repeat homology
C;Keywords: duplication
Query Match 64.7%; Score 44; DB 2; Length 33;
Best Local Similarity 54.5%; Pred. No. 0.28; Mismatches 2; Indels 0; Gaps 0;
Matches 6; Conservative 3;
Qy 1 CFAWQAMRKV 11
Db 19 CYQWQRMKRL 29

RESULT 4
G71633
ADP, ATP carrier protein (tlcs) RP739 - Rickettsia prowazekii
C;Species: Rickettsia prowazekii
C;Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
C;Accession: G71633
R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichert-Ponten, T.; Alsmark,
Nature 396, 133-140, 1998
A;Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A;Reference number: A71630; MUID:99039499; PMID:9823893
A;Accession: G71633
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-500 <AND>
A;Cross-references: GB:AJ235273; GB:AJ235269; NID:93861237; PIDN:CAA15167.1; PID:93861;
A;Experimental source: strain Madrid E
C;Genetics:
A;Gene: tlc5; RP739
C;Superfamily: rickettsial-type ATP/ADP translocase
Query Match 63.2%; Score 43; DB 2; Length 500;
Best Local Similarity 63.6%; Pred. No. 5.8; Mismatches 2; Indels 0; Gaps 0;
Matches 7; Conservative 2;
Qy 1 CFAWQAMRKV 11
Db 482 CFAWIYAVRKI 492

RESULT 5
AH3147
hypothetical protein Atu4804 [imported] - Agrobacterium tumefaciens (strain C58, Dupont
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C;Accession: AH3147
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuyavin, T.; Levy, R.; Li, M.; McCle:

```

J. Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Rao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, S.E.W.

A>Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AH3147

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-121 <KUR>

A:Cross-references: GB:AB008689; PIDN:AAU45598.1; PID:gl7743317; GSPDB:GN00187

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu4804

A:Map position: linear chromosome

Query Match 61.8%; Score 42; DB 2; Length 121;

Best Local Similarity 63.6%; Pred. No. 2.2;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFAWORAMRKV 11

DB 14 CLAWQRNRV 24

RESULT 6

A28438

lactoferrin precursor - mouse

N:Alternate names: lactotransferrin

C:Species: Mus musculus (house mouse)

C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

A:Accession: A28438; A1205

R:Parent: B.T.; Teng, C.T.

J. Biol. Chem. 262, 10134-10139, 1987

A>Title: Lactotransferrin is the major estrogen inducible protein of mouse uterine secret

A:Reference number: A25596; MUID:87280033; PMID:3611056

A:Accession: A28438

A:Molecule type: mRNA

A:Residues: 3-707 <PEN>

A:Cross-references: EMBL:J03298

R:Li, Y.; Teng, C.T.

J. Biol. Chem. 266, 21880-21885, 1991

A>Title: Characterization of estrogen-responsive mouse lactoferrin promoter.

A:Reference number: A1205; MUID:92042099; PMID:1939212

A:Accession: A1205

A:Molecule type: DNA

A:Residues: 1-15 <LIU>

A:Cross-references: GB:M74778

C:Superfamily: transferrin; transferrin repeat homology

F:1-19/Domain: duplication; glycoprotein

F:20-707/Product: lactotransferrin #status predicted <SIG>

F:358-695/Domain: transferrin repeat homology <TRH2>

F:494/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 61.8%; Score 42; DB 1; Length 707;

Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFAWORAMRKV 11

DB 37 CLRWQNRKV 47

RESULT 7

D33876

carcinoembryonic antigen homolog 4 - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 27-Feb-1990 #sequence\_revision 27-Feb-1990 #text\_change 23-Jul-1999

C:Accession: D33876

R:Kodelia, V.; Lucas, K.; Barnert, S.; von Kleist, S.; Thompson, J.A.; Zimmermann, W.

J. Biol. Chem. 264, 6906-6912, 1989

A>Title: Identification of a carcinoembryonic antigen gene family in the rat. Analysis d

A:Reference number: A33876; MUID:89214106; PMID:2708349

A:Accession: D33876

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-114 <KOD>

A:Cross-references: GB:M60026; NID:G203407; PIDN:AAA40911.1; PID:G554427; GB:J04626; C

C:Superfamily: carcinoembryonic antigen; carcinoembryonic precursor amino-term

Query Match 58.8%; Score 40; DB 2; Length 114;

Best Local Similarity 60.0%; Pred. No. 4.8;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 FAWORAMEKV 11

DB 42 FAWYRGLRKI 51

RESULT 8

A39654

cell cycle arrest protein BUB2 - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein YMR796.08c; protein YMR055c

C:Species: Saccharomyces cerevisiae

C>Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 21-Jul-2000

A:Accession: A39654; S54555; S19034

R:Hoyt, M.A.; Totis, L.; Roberts, B.T.

Cell 66, 507-517, 1991

A>Title: Saccharomyces cerevisiae genes required for cell cycle arrest in response to

A:Reference number: A39654; MUID:91330299; PMID:1651171

A:Accession: A39654

A:Molecule type: DNA

A:Residues: 1-306 <HOY>

A:Cross-references: GB:M64706; NID:gl71133; PIDN:AAA16885.1; PID:gl71135

R:Devlin, K.; Churcher, C.M.

submitted to the EMBL Data Library, May 1995

A:Reference number: S54548

A:Accession: S54555

A:Molecule type: DNA

A:Residues: 1-306 <DEV>

A:Cross-references: EMBL:Z49703; NID:gl71880; PIDN:CAA89765.1; PID:gl71888; GSPDB:GN00

C:Genetics:

A:Gene: SGD:BUB2; MIPS:YMR055C

A:Cross-references: SGD:S0004659; MIPS:YMR055C

A:Map position: 13R

C:Superfamily: cell cycle arrest protein BUB2

C:Keywords: transmembrane protein

F:232-248/Domain: transmembrane #status predicted <TMW>

Query Match 58.8%; Score 40; DB 1; Length 306;

Best Local Similarity 60.0%; Pred. No. 12;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFAWORAMRK 10

DB 108 CFAWQTQQR 117

RESULT 9

AB0858

hypothetical protein STV3070 [imported] - Salmonella enterica subsp. enterica serovar

C:Species: Salmonella enterica subsp. enterica serovar Typhi

A>Note: this species has also been called Salmonella typhi

C>Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 09-Nov-2001

C:Accession: AB0858

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church

th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farra

, S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K

A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica ser.

A:Reference number: AB0502; PMID:11677608

A:Accession: AB0858

A>Status: preliminary

A:Molecule type: DNA

QY 1 CFAWQAMRKVR 12  
 |||:  
 Db 8 CFTWEYARHVR 19  
 |||:  
 RESULT 12  
 E82395 methyl-accepting chemotaxis protein VCA0974 [imported] - Vibrio cholerae (strain N169)  
 C:Species: Vibrio cholerae  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
 C:Accession: E82395  
 R:Reidberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.;  
 charadson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Drsgol, I.; Sellers  
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A:Reference number: A82035; MUID:20406833; PMID:10952301  
 A:Accession: E82395  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-561 <HEI>  
 A:Cross-references: GB:AE004423; GB:AE003853; NID:9658400; PIDN:AAF96870.1; GSPDB:GN  
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C:Genetics:  
 A:Gene: VCA0974  
 A:Map position: 2  
 Query Match 55.9%; Score 38; DB 2; Length 561;  
 Best Local Similarity 63.6%; Pred. NO. 51;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 FAWQAMRKVR 12  
 |||:  
 Db 7 FAWQALRLKR 17  
 |||:  
 RESULT 13  
 AC3242 NTP pyrophosphohydrolase, MutT family [imported] - Agrobacterium tumefaciens (strain  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
 C:Accession: AC3242  
 R:Wood, D.W.; Sebubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo  
 erage, G.; Giller, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCl  
 i Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam  
 ster, E.W.  
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A:Reference number: AB2577; PMID:11743193  
 A:Accession: AC3242  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-161 <KUR>  
 A:Cross-references: GB:AE008630; PIDN:RAL46353.1; PID:gl7744142; GSPDB:GN00189  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: tiorf74  
 A:Genome: plasmid  
 Query Match 54.4%; Score 37; DB 2; Length 161;  
 Best Local Similarity 60.0%; Pred. NO. 23;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 AWQAMRKVR 12  
 |||:  
 Db 51 AWEALREVR 60  
 |||:



C>Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 05-Nov-1999  
 C/Accession: C71305  
 R/Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin  
 rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo  
 they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
 Science 281, 375-388, 1998  
 A>Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
 A/Reference number: A71250; MUID:9832770; PMID:9665876  
 A/Accession: C71305  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-202 <COL>  
 A/Cross-references: GB:AE001234; GB:AE000520; NID:g3322881; PIDN:AAC65572.1; PID:g332289  
 A/Experimental source: strain Nichols  
 C/Genetics:  
 A/Gene: TP0594

Query Match 54.4%; Score 37; DB 2; Length 202;  
 Best Local Similarity 87.5%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 WQAMRKV 11  
 DB 21 WQAMRKV 28

## RESULT 15

H97451  
 Pyridoxamine 5'-phosphate oxidase (A9179611) [imported] - Agrobacterium tumefaciens (str  
 C/Species: Agrobacterium tumefaciens  
 C/Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002  
 C/Accession: H97451  
 R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,  
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
 Science 294, 2323-2328, 2001  
 A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
 A/Reference number: A97359; PMID:11743194  
 A/Accession: H97451  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-206 <KUR>  
 A/Cross-references: GB:AE007869; PIDN:AAK86569.1; PID:g15155733; GSPDB:GN00169  
 C/Genetics:  
 A/Gene: AGR\_C\_1381  
 A/Map position: circular chromosome  
 C/Superfamily: pyridoxamine-phosphate oxidase

Query Match 54.4%; Score 37; DB 2; Length 206;  
 Best Local Similarity 50.0%; Pred. No. 29;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFAMQAMRKVR 12  
 DB 88 CFHMKSLRRQVR 99

Search completed: February 21, 2003, 08:02:49  
 Job time : 11.6047 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:28:06 ; Search.time 5.2093 Seconds  
(without alignments)  
95.544 Million cell updates/sec

Title: US-09-743-107b-95

Perfect score: 68

Sequence: 1 CFANQRMKVR 12

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	83.8	711	1 TRFL_HUMAN	P02788 homo sapien
2	47	69.1	708	1 TRFL_CAMDR	Q9tun0 camelus dro
3	47	69.1	708	1 TRFL_CAPHI	Q29477 capra hircu
4	43	63.2	500	1 TLCE_RICPR	O05962 rickettsia
5	42	61.8	707	1 TRFL_MOUSE	P08071 mus musculus
6	40	58.8	306	1 BUB2_YEAST	P26448 saccharomyc
7	39	57.4	1135	1 PHYC_SORBI	P93528 sorghum bic
8	37	54.4	275	1 IL2A_BOVIN	P12342 bos taurus
9	37	54.4	275	1 IL2A_SHEEP	P26898 ovina aries
10	37	54.4	303	1 UL24_HSVSA	Q01005 herpesvirus
11	37	54.4	601	1 WD66_PHYPO	P90587 physarum po
12	37	54.4	1137	1 PHYC_ORYZA	Q92w19 oryza sativ
13	37	54.4	1179	1 ATX1_ARATH	O91t02 arabidopsis
14	37	54.4	2273	1 ABCR_HUMAN	P78363 homo sapien
15	37	54.4	4568	1 DYHE_CHLRE	Q39865 chlamydomon
16	36	52.9	267	1 RS3A_DROME	F55830 drosophila
17	36	52.9	363	1 CTNS_DROME	Q9vcr7 drosophila
18	36	52.9	695	1 TRFL_HORSE	P77811 equus cabal
19	36	52.9	805	1 L100_ADR02	P24932 human adeno
20	36	52.9	807	1 L100_ADR05	P24933 human adeno
21	36	52.9	956	1 SVV_TREPA	O83998 treponema p
22	36	52.9	1208	1 RCQ4_HUMAN	Q94761 homo sapien
23	35	51.5	242	1 PYG_GALSU	Q02074 gallieria s
24	35	51.5	275	1 VA16_VACCV	P16710 vaccinia vi
25	35	51.5	378	1 VA16_VACCC	P20993 vaccinia vi
26	35	51.5	538	1 R060_HUMAN	P10155 homo sapien
27	35	51.5	538	1 R060_MOUSE	O08848 mus musculus
28	35	51.5	538	1 R060_XENLA	P42700 xenopus lae
29	35	51.5	1214	1 TSGA_RAT	Q63679 rattus norv
30	35	51.5	1574	1 RPOC_AQUAE	O67763 aquifex aeo
31	35	51.5	1576	1 RPOC_AQUAP	Q9x6y2 aquifex pyr
32	34	50.0	85	1 PMRD_SALTY	P37589 salmonella
33	34	50.0	116	1 MERT_SALTI	P04336 salmonella

34 34 50.0 234 1 PYG\_CVACA O19899 cyanidium c  
35 34 50.0 235 1 FL3L\_HUMAN P49771 homo sapien  
36 34 50.0 268 1 YC59\_MYCTU Q11059 mycobacteri  
37 34 50.0 270 1 PDXH\_MYXXA P21159 myxococcus  
38 34 50.0 278 1 XRC2\_MOUSE Q9CX47 mus musculus  
39 34 50.0 281 1 VNSI\_INELE P03502 influenza b  
40 34 50.0 385 1 GP68\_HUMAN Q15743 homo sapien  
41 34 50.0 435 1 DHOM\_METGL P37143 methylobacti  
42 34 50.0 1024 1 CARB\_PYRAE Q82Y48 pyrobaculum  
43 34 50.0 1131 1 PHY\_PINSY Q41046 pinus sylve  
44 34 50.0 1293 1 XPC\_DROME Q24595 drosophila  
45 33 48.5 129 1 VP70\_MYCTU Q50651 mycobacteri

## ALIGNMENTS

RESULT 1  
TRFL\_HUMAN STANDARD; PRT: 711 AA.  
AC P02788; Q16780; Q16785; Q16786; Q00756; Q9HLZ3; Q96KZ4;  
AC Q96KZ5;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Lactoferrin precursor (Lactoferrin) (Contains: Lactoferrin A;  
DE Lactoferrin B; Lactoferrin C).  
GN LIF OR LF.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary Gland;  
RX MEDLINE=90384839; PubMed=2402455;  
RA Rey M.W., Woloshuk S.L., de Boer H.A., Pieper F.R.;  
RT "Complete nucleotide sequence of human mammary gland lactoferrin.";  
RL Nucleic Acids Res. 18:5288-5288(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary Gland;  
RA Cho Y.Y.;  
RL Thesis (1994), Genetic Engineering Research Institute / Taejon, Korea.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Connely O.M.;  
RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary Gland;  
RA Liang Q., Jimenez-Flores R., Richardson T.;  
RL "Molecular cloning and sequence analysis of human lactoferrin.";  
RN Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Bone marrow;  
RA Wei X., Han J., Rao T.A.;  
RL "Human neutrophil lactoferrin coding and 5' flanking region DNA  
sequences.";  
RN Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RA Cheng H., Chen X., Huan L.;  
RL "cDNA cloning and sequence analysis of human lactoferrin.";  
RN Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Prostate;  
RA Strausberg R.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
RN [8]

RC SEQUENCE OF 3-711 FROM N.A.  
RC TISSUE-Mammary gland;  
RX MEDLINE=90326549; PubMed=2374734;  
RA Powell M.J., Ogden J.E.;  
RT "Nucleotide sequence of human lactoferrin cDNA.";  
RL Nucleic Acids Res. 18:4013-4013(1990).  
RN [9]  
RP SEQUENCE OF 20-711  
RX MEDLINE=9076667; PubMed=6510420;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Schoentgen F.,  
RA Legrand D., Spik G., Montreuil J., Jolles P.;  
RT "Human lactotransferrin: amino acid sequence and structural  
RT comparisons with other transferrins.";  
RL Eur. J. Biochem. 145:659-666(1984).  
RN [10]  
RP PRELIMINARY SEQUENCE OF 20-73; 134-171; 257-278 AND 347-530.  
RX MEDLINE=82046817; PubMed=6794640;  
RA Metz-Boutigue M.-H., Mazurier J., Jolles J., Spik G., Montreuil J.,  
RA Jolles P.;  
RT "The present state of the human lactotransferrin sequence. Study and  
RT alignment of the cyanogen bromide fragments and characterization of  
RT N- and C-terminal domains.";  
RL Biochim. Biophys. Acta 670:243-254(1981).  
RN [11]  
RP SEQUENCE OF 609-711.  
RX MEDLINE=82262043; PubMed=7049727;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,  
RA Jolles P.;  
RT "An 88 amino acid long C-terminal sequence of human  
RT lactotransferrin.";  
RL FEBS Lett. 142:1107-110(1982).  
RN [12]  
RP SEQUENCE OF 436-711 FROM N.A.  
RX MEDLINE=88001031; PubMed=3477300;  
RA Rado T.A., Wei X., Benz E.J. Jr.;  
RT "Isolation of lactoferrin cDNA from a human myeloid library and  
RT expression of mRNA during normal and leukemic myelopoiesis.";  
RL Blood 70:989-993(1987).  
RN [13]  
RP SEQUENCE OF 237-711 FROM N.A.  
RA McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,  
RA Nhan M., Parnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,  
RA Choj L., la Bastide M., Kaplan N., Greco T., Touchman J., Muzny D.,  
RA Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M., Porcel B.M.,  
RA Dragan Y., Giacalone J., Pae A., Powell E., Solinsky K.A., Desilva U.,  
RA Diaz-Perez S., Zhou X., Yu Y., Watanabe M., Doggett N., Garcia D.,  
RA Sagripanti J.L.;  
RT Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
RN [14]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND REVISIONS.  
RX MEDLINE=90064528; PubMed=2585506;  
RA Anderson B.F., Baker H.M., Norris G.E., Rice D.W., Baker E.N.;  
RT "Structure of human lactoferrin: crystallographic structure analysis  
RT and refinement at 2.8-A resolution.";  
RL J. Mol. Biol. 209:711-734(1989).  
RN [15]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RA Haridas M., Anderson B.F., Baker E.N.;  
RT "Structure of human diferric lactoferrin refined at 2.2-A  
RT resolution.";  
RL Acta Crystallogr. D 51:629-646(1995).  
RN [16]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 20-353.  
RX MEDLINE=97156796; PubMed=9003186;  
RA Nicholson H., Anderson B.F., Bland T., Shewry S.C., Tweedie J.W.,  
RA Baker E.N.;  
RT "Mutagenesis of the histidine ligand in human lactoferrin: iron  
RT binding properties and crystal structure of the histidine-  
RT 253-->methionine mutant.";  
RL Biochemistry 36:341-346(1997).  
RN [17]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX MEDLINE=99190892; PubMed=10089347;  
RA Sun X.L., Baker H.M., Shewry S.C., Jameson G.B., Baker E.N.;  
RT "Structure of recombinant human lactoferrin expressed in *Aspergillus*  
RT *awamori*.";  
RL Acta Crystallogr. D 55:403-407(1999).  
RN [18]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=99192677; PubMed=10089508;  
RA Jameson G.B., Anderson B.F., Norris G.E., Thomas D.H., Baker E.N.;  
RT "Structure of human apolactoferrin at 2.0-A resolution. Refinement  
RT and analysis of ligand-induced conformational change.";  
RL Acta Crystallogr. D 54:1319-1335(1998).  
RN [19]  
RP CHARACTERIZATION OF LACTOFERROXINS.  
RX MEDLINE=91166929; PubMed=1369293;  
RA Tani F., Iio K., Chiba H., Yoshikawa M.;  
RT "Isolation and characterization of opioid antagonist peptides derived  
RT from human lactoferrin.";  
RL Agric. Biol. Chem. 54:1803-1810(1990).  
RN [20]  
RP VARIANTS THR-30 AND ARG-48.  
RX PubMed=9873069;  
RA Klintworth G.K., Sommer J.R., Obrian G., Han L., Ahmed M.N.,  
RA Qumsiyeh M.B., Lin P.-Y., Basti S., Reddy M.K., Kanai A., Hotta Y.,  
RA Sugar J., Kumaramanickavel G., Munier F., Schorderet D.F.,  
RA El Matiri L., Iwata F., Kaiser-Kupfer M., Nagata M., Nakayasu K.,  
RA Hejtmancik J.F., Teng C.T.;  
RT "Familial subepithelial corneal amyloidosis (gelatinous drop-like  
RT corneal dystrophy): exclusion of linkage to lactoferrin gene.";  
RL Mol. Vision 4:31-32(1998).  
RN [21]  
RP FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
CC OF AN ANION, USUALLY BICARBONATE.  
CC -1- FUNCTION: LACTOFERROXIN A, B AND C HAVE OPIOID ANTAGONIST  
CC ACTIVITY. LACTOFERROXIN A SHOWS PREFERENCE FOR MU-RECEPTORS, WHILE  
CC LACTOFERROXIN B AND C HAVE SOMEWHAT HIGHER DEGREES OF PREFERENCE  
CC FOR KAPPA-RECEPTORS THAN FOR MU-RECEPTORS.  
CC -1- SUBUNIT: MONOMER.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X53961; CAA37914.1; -  
CC EMBL; U07643; AAB60324.1; -  
CC EMBL; M93150; AAA36159.1; -  
CC EMBL; M83202; AAA59511.1; -  
CC EMBL; M83205; AAA58656.1; -  
CC EMBL; M18642; AAA86665.1; -  
CC EMBL; AF332168; AAG48753.1; -  
CC EMBL; SC015822; AAHL5822.1; -  
CC EMBL; SC015823; AAHL5823.1; -  
CC EMBL; M73700; AAA59479.1; -  
CC EMBL; X52941; CAA37116.1; -  
CC EMBL; U95626; AAB57795.1; -  
CC PIR; S11228; TFHUL.  
CC PDB; 1LCF; 31-AUG-94.  
CC PDB; 1LCT; 31-OCT-93.  
CC PDB; 1LFG; 31-JUL-94.  
CC PDB; 1LFH; 31-OCT-93.  
CC PDB; 1LFI; 31-OCT-93.  
CC PDB; 1LGB; 31-AUG-94.  
CC PDB; 1LGC; 31-AUG-94.  
CC PDB; 1BKA; 08-NOV-96.  
CC PDB; 1DSN; 08-MAR-96.  
CC PDB; 1HSE; 12-MAR-97.  
CC PDB; 1VFD; 21-APR-97.

Query Match 83.8%; Score 57; DB 1; Length 711;  
 Best Local Similarity 83.3%; Pred. No. 0.0052;  
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12  
 |||||  
 DB 39 CFQWQRMKVR 50

RESULT 2  
 TRFL\_CAMDR STANDARD; PRT; 708 AA.

AC Q9TUM0; Q9WZS5;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin).  
 GN LTF.  
 OS Camelus dromedarius (Dromedary) (Arabian camel).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.  
 OX NCBI\_TaxID=9898;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Somali; TISSUE=Lactating mammary gland;  
 RA Kappeler S.R., Ackermann M., Farah Z., Puhon Z.;  
 RT "Sequence analysis of camel (Camelus dromedarius) lactoferrin.";  
 RL Int. Dairy J. 9:481-486(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Paramasivam M., Srinivasan A., Singh R., Sahani M.S., Singh T.P.;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 OF AN ANION, USUALLY BICARBONATE (BY SIMILARITY).  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AJ131674; CAB53387.1; -  
 DR EMBL; AF165879; AAF82241.1; -  
 DR HSSP; O77811; 1B1X.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 2.  
 DR PRINTS; PR00422; transferrin.  
 DR SMART; SM00094; TR\_FER\_2  
 DR PROSITE; PS00205; TRANSFERRIN 1; 2.  
 DR PROSITE; PS00206; TRANSFERRIN 2; 2.  
 DR PROSITE; PS00207; TRANSFERRIN 3; 2.  
 KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 KW Signal.  
 FT SIGNAL 1 19 BY SIMILARITY.  
 FT CHAIN 20 708 LACTOTRANSFERRIN.  
 FT REPEAT 20 363 1.  
 FT REPEAT 364 708 2.  
 FT DISULFID 28 64 BY SIMILARITY.  
 FT DISULFID 38 55 BY SIMILARITY.  
 FT DISULFID 134 217 BY SIMILARITY.  
 FT DISULFID 176 192 BY SIMILARITY.  
 FT DISULFID 189 200 BY SIMILARITY.  
 FT DISULFID 250 264 BY SIMILARITY.  
 FT DISULFID 367 399 BY SIMILARITY.

FT DISULFID 377 390 BY SIMILARITY.  
 FT DISULFID 424 703 BY SIMILARITY.  
 FT DISULFID 444 666 BY SIMILARITY.  
 FT DISULFID 476 551 BY SIMILARITY.  
 FT DISULFID 500 694 BY SIMILARITY.  
 FT DISULFID 510 524 BY SIMILARITY.  
 FT DISULFID 521 534 BY SIMILARITY.  
 FT DISULFID 592 606 BY SIMILARITY.  
 FT DISULFID 644 649 BY SIMILARITY.  
 FT METAL 79  
 FT METAL 111  
 FT METAL 211  
 FT METAL 272  
 FT METAL 414  
 FT METAL 452  
 FT METAL 545  
 FT METAL 614  
 FT BINDING 140  
 FT BINDING 140  
 FT BINDING 482  
 FT CARBOHYD 252  
 FT CARBOHYD 385  
 FT CARBOHYD 537  
 FT CARBOHYD 594  
 FT CONFLICT 261  
 FT CONFLICT 304  
 FT CONFLICT 330  
 FT CONFLICT 492  
 FT CONFLICT 506  
 FT CONFLICT 609  
 FT CONFLICT 642  
 SQ SEQUENCE 708 AA; 77211 MW; 0B0C175A0B69D430 CRC64;

Query Match 69.1%; Score 47; DB 1; Length 708;  
 Best Local Similarity 66.7%; Pred. No. 0.37;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQWQRMKVR 12  
 |||||  
 DB 38 CAQWQRMKVR 49

RESULT 3  
 TRFL\_CAPHI STANDARD; PRT; 708 AA.

ID TRFL\_CAPHI  
 AC Q29477; Q29479;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (lactoferrin).  
 GN LTF.  
 OS Capra hircus (Goat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Capra.  
 OX NCBI\_TaxID=9925;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Lee T., Yu S., Kim S., Lee K., Yu D.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RX MEDLINE=94380047; PubMed=8093048;  
 RA le Provost F., Nocard M., Guerin G., Martin P.;  
 RT "Characterization of the goat lactoferrin cDNA. Assignment of the  
 relevant locus to bovine U12 syntenic group.";  
 RL Biochem. Biophys. Res. Commun. 203:1324-1332(1994).  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 OF AN ANION, USUALLY BICARBONATE.  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Secreted.

CC -I- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
CC -I- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X53857; AAA97958.1; -  
DR EMBL; X78902; CAA55517.1; -  
DR HSP; O77698; 1CE2.  
DR InterPro; IPR001156; Transferrin.  
DR Pfam; PF00405; transferrin; 2.  
DR PRINTS; PR00422; TRANSFERRIN.  
DR SMART; SM00094; TF\_FER; 2.  
DR PROSITE; PS00205; TRANSFERRIN\_1; 2.  
DR PROSITE; PS00206; TRANSFERRIN\_2; 2.  
DR PROSITE; PS00207; TRANSFERRIN\_3; 2.  
KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
KW Signal.  
FT SIGNAL 1 19 BY SIMILARITY.  
FT CHAIN 20 708 LACTOTRANSFERRIN.  
FT REPEAT 20 363 1.  
FT REPEAT 364 708 2.  
FT DISULFID 38 64 BY SIMILARITY.  
FT DISULFID 28 55 BY SIMILARITY.  
FT DISULFID 134 217 BY SIMILARITY.  
FT DISULFID 176 192 BY SIMILARITY.  
FT DISULFID 189 200 BY SIMILARITY.  
FT DISULFID 250 264 BY SIMILARITY.  
FT DISULFID 367 399 BY SIMILARITY.  
FT DISULFID 377 390 BY SIMILARITY.  
FT DISULFID 424 703 BY SIMILARITY.  
FT DISULFID 444 666 BY SIMILARITY.  
FT DISULFID 476 551 BY SIMILARITY.  
FT DISULFID 500 694 BY SIMILARITY.  
FT DISULFID 510 524 BY SIMILARITY.  
FT DISULFID 521 534 BY SIMILARITY.  
FT DISULFID 592 606 BY SIMILARITY.  
FT DISULFID 644 649 BY SIMILARITY.  
FT METAL 79 79 IRON 1 (BY SIMILARITY).  
FT METAL 111 111 IRON 1 (BY SIMILARITY).  
FT METAL 211 211 IRON 1 (BY SIMILARITY).  
FT METAL 272 272 IRON 2 (BY SIMILARITY).  
FT METAL 414 414 IRON 2 (BY SIMILARITY).  
FT METAL 452 452 IRON 2 (BY SIMILARITY).  
FT METAL 545 545 IRON 2 (BY SIMILARITY).  
FT METAL 614 614 IRON 2 (BY SIMILARITY).  
FT BINDING 140 140 ANION (BY SIMILARITY).  
FT BINDING 482 482 ANION (BY SIMILARITY).  
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 495 495 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 564 564 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 56 56 I -> V (IN REF. 2).  
FT CONFLICT 88 88 L -> R (IN REF. 2).  
FT CONFLICT 124 124 F -> K (IN REF. 2).  
FT CONFLICT 154 154 F -> P (IN REF. 2).  
FT CONFLICT 304 304 S -> R (IN REF. 2).  
FT CONFLICT 414 414 D -> G (IN REF. 2).  
SQ SEQUENCE 708 AA; 77358 MW; F2EDA3C83539960D CRC64;

Query Match 69.1%; Score 47; DB 1; Length 708;  
Best Local Similarity 63.6%; Pred. No. 0.37;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFAWQRMKRV 11  
| : | | | :  
DB 38 CYQWQRMKRL 48

RESULT 4  
TLCE\_RICPR STANDARD; PRT; 500 AA.  
AC C05962;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE ADP-ATP carrier protein 5 (ADP/ATP translocase 5).  
GN TLCE OR TLCS OR RP739.  
OS Rickettsia prowazekii.  
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;  
OC Rickettsiaceae; Rickettsia.  
OX NCBI\_TaxID=782;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Madrid E;  
RX MEDLINE=99039499; PubMed=9823893;  
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,  
RA Sacheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,  
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;  
RT "The genome sequence of Rickettsia prowazekii and the origin of  
RT mitochondria.";  
RL Nature 396:133-140(1998).  
RN [2]  
RP SEQUENCE OF 325-500 FROM N.A.  
RC STRAIN=Madrid E;  
RX MEDLINE=97419517; PubMed=9274032;  
RA Andersson J.O., Andersson S.G.E.,  
RT "Genomic rearrangements during evolution of the obligate  
RT intracellular parasite Rickettsia prowazekii as inferred from an  
RT analysis of 52015 bp nucleotide sequence.";  
RL Microbiology 143:2783-2795(1997).  
CC -I- FUNCTION: PROVIDES THE RICKETTSIAL CELL WITH HOST ATP IN  
CC EXCHANGE FOR RICKETTSIAL ADP. THIS IS AN OBLIGATE EXCHANGE SYSTEM.  
CC THIS ENERGY ACQUIRING ACTIVITY IS AN IMPORTANT COMPONENT OF  
CC RICKETTSIAL PARASITISM (BY SIMILARITY).  
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -I- SIMILARITY: BELONGS TO THE ADP/ATP TRANSLOCASE TLC FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AJ235273; CAA15167.1; -  
DR EMBL; Y11778; CAA72457.1; -  
DR InterPro; IPR004667; ADP\_ATP\_car.  
DR Pfam; PF03219; TLC; 1.  
DR TIGRFAMs; TIGR00769; AAA; 1.  
KW Transmembrane; Transport; ATP-binding; Multigene family;  
KW Complete proteome.  
FT TRANSMEM 26 46 POTENTIAL.  
FT TRANSMEM 52 82 POTENTIAL.  
FT TRANSMEM 94 114 POTENTIAL.  
FT TRANSMEM 149 169 POTENTIAL.  
FT TRANSMEM 184 204 POTENTIAL.  
FT TRANSMEM 224 244 POTENTIAL.  
FT TRANSMEM 287 307 POTENTIAL.  
FT TRANSMEM 328 348 POTENTIAL.  
FT TRANSMEM 357 377 POTENTIAL.  
FT TRANSMEM 381 401 POTENTIAL.  
FT TRANSMEM 469 489 POTENTIAL.  
SQ SEQUENCE 500 AA; 57073 MW; FE3DB48D08CF5F72 CRC64;  
Query Match 63.2%; Score 43; DB 1; Length 500;  
Best Local Similarity 63.6%; Pred. No. 1.4;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 CFAMQAMRKV 11
DB 482 CFAMQAMRKV 492

RESULT 5
TRFL_MOUSE
ID TRFL_MOUSE STANDARD; PRT; 707 AA.
AC P08071; P70690; Q61799; Q9222P2;
DT 01-AUG-1988 (Rel. 08, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lactotransferrin precursor (Lactoferrin).
GN LTF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RX MEDLINE=87280033; PubMed=3611056;
RA Pentecost B.T.; Teng C.T.;
RT "Lactotransferrin is the major estrogen inducible protein of mouse
RT uterine secretions.";
RL J. Biol. Chem. 262:10134-10139 (1987).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=uterus;
RA Morlishi K.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Strauberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 1-14 FROM N.A.
RX MEDLINE=92042099; PubMed=1939212;
RA Liu Y.; Teng C.T.;
RT "Characterization of estrogen-responsive mouse lactoferrin promoter.";
RL J. Biol. Chem. 266:21880-21885 (1991).
CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
CC OF AN ANION, USUALLY BICARBONATE.
CC -!- SUBUNIT: MONOMER.
CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J03298; AAA40525.1; -
CC EMBL; D88510; BAA13633.1; -
CC EMBL; BC006904; AAA06904.1; -
CC EMBL; M74778; AAA39427.1; -
CC PIR; A28438; A28438.
CC HSP; P02788; 1CE6.
CC MGD; MGI:96837; Ltf.
CC InterPro; IPR001156; Transferrin.
CC Pfam; PF00405; transferrin; 2.
CC PRINTS; PR00422; TRANSFERRIN.
CC SMART; SM00094; TR_FER; 2.
CC PROSITE; PS00205; TRANSFERRIN_1; 1.
CC PROSITE; PS00206; TRANSFERRIN_2; 2.
CC PROSITE; PS00207; TRANSFERRIN_3; 2.
KW Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;
KW Signal.

FT SIGNAL 1 19
FT CHAIN 20 707
FT REPEAT 20 357
FT REPEAT 358 707
FT DISULFID 27 63
FT DISULFID 37 54
FT DISULFID 133 216
FT DISULFID 175 191
FT DISULFID 188 199
FT DISULFID 249 263
FT DISULFID 366 398
FT DISULFID 376 389
FT DISULFID 423 702
FT DISULFID 443 665
FT DISULFID 475 550
FT DISULFID 499 693
FT DISULFID 509 523
FT DISULFID 520 533
FT DISULFID 591 605
FT DISULFID 643 648
FT METAL 78
FT METAL 110
FT METAL 210
FT METAL 271
FT METAL 413
FT METAL 451
FT METAL 454
FT METAL 613
FT METAL 613
FT BINDING 139
FT BINDING 481
FT CARBOHYD 118
FT CARBOHYD 494
FT CONFLICT 1
FT CONFLICT 25
FT CONFLICT 82
FT CONFLICT 359
FT CONFLICT 382
FT CONFLICT 449
FT CONFLICT 629
SQ SEQUENCE 707 AA; 77865 MW; P26AE0340A4C19A8 CRC64;

Query Match 61.8%; Score 42; DB 1; Length 707;
Best Local Similarity 63.6%; Pred. No. 3.1;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CFAMQAMRKV 11
DB 37 CLRQWNEVRKV 47

RESULT 6
BUB2 YEAST STANDARD; PRT; 306 AA.
ID BUB2 YEAST STANDARD; PRT; 306 AA.
AC P26438;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mitotic check point protein BUB2 (Cell cycle arrest protein BUB2).
GN BUB2 OR YMR055C OR YMR796.08C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=91330299; PubMed=1651171;
RA Hoyt M.A.; Totis L.; Roberts B.T.;
RT "S. cerevisiae genes required for cell cycle arrest in response to
RT loss of microtubule function.";
RL Cell 66:507-517 (1991).
RN [2]
RP SEQUENCE FROM N.A.

```

RC STRAIN=9288C / AB972;  
RA Devlin K., Churcher C.M., Barrall B.G., Rajandream M.A.;  
RL Submitted (MAY-1995) to the EMBL/GenBank/DBSJ databases.  
RN [3]  
RP CHARACTERIZATION.  
RX MEDLINE=21385309; PubMed=11493673;  
RA Lee S.E., Jensen S., Frenz L.M., Johnson A.L., Fesquet D.,  
RA Johnston L.H.;  
RA "the Bub2-dependent mitotic pathway in yeast acts every cell cycle and  
RT regulates cytokinesis.";  
RL J. Cell Sci. 114:2345-2354 (2001).  
CC -!- FUNCTION: Part of a checkpoint which monitors spindle integrity  
CC and prevents premature exit from mitosis. This cell-cycle arrest  
CC depends upon inhibition of the G-protein Tem1 by the BFA1/BUB2  
CC complex.  
CC -!- SUBUNIT: Interacts with BFA1.  
CC -!- SUBCELLULAR LOCATION: Spindle poles.  
CC -!- SIMILARITY: TO S.POMBE CDC16.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M64706; AAA16885.1; -;  
CC EMBL; Z49703; CAA89765.1; -;  
CC PIR; A39654; A39654.  
CC SGD; S0004659; BUB2.  
CC InterPro; IPR000195; RabGAP\_TBC.  
CC Pfam; PF00566; TBC; 1.  
CC SMART; SM00164; TBC; 1.  
CC Cell cycle; Mitosis.  
CC SEQUENCE 306 AA; 35027 MW; A1DDBP548E81EA3 CRC64;  
  
Query Match 58.8%; Score 40; DB 1; Length 306;  
Best Local Similarity 60.0%; Pred. No. 3.1;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 CFAWQAMRK 10  
DB 108 CFAWQTOQR 117  
  
RESULT 7  
ID PHYC\_SORBI STANDARD; PRT; 1135 AA.  
AC P9328;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Phytochrome C.  
DE PHYC.  
OS Sorghum bicolor (Sorghum) (Sorghum vulgare).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Panicoideae; Andropogoneae; Sorghum.  
OX NCBI\_TaxID=4558;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97198556; PubMed=9046599;  
RA Childs K.L., Miller F.R., Cordonnier-Pratt M.-M., Pratt L.H.,  
RA Morgan P.W., Muller J.E.;  
RA "The Sorghum bicolor photoperiod sensitivity gene, Ma3, encodes a  
RT phytochrome B.";  
RL Plant Physiol. 113:611-619 (1997).  
CC -!- FUNCTION: REGULATORY PHOTORECEPTOR WHICH EXISTS IN TWO FORMS THAT  
CC ARE REVERSIBLY INTERCONVERTIBLE BY LIGHT. THE PR FORM THAT ABSORBS  
CC MAXIMALLY IN THE RED REGION OF THE SPECTRUM AND THE PFR FORM THAT  
CC ABSORBS MAXIMALLY IN THE FAR-RED REGION. PHOTOCONVERSION OF PR IN  
CC PFR INDUCES AN ARRAY OF MORPHOGENIC RESPONSES, WHEREAS

CC RECONVERSION OF PFR TO PR CANCELS THE INDUCTION OF THOSE  
CC RESPONSES. PFR CONTROLS THE EXPRESSION OF A NUMBER OF NUCLEAR  
CC GENES INCLUDING THOSE ENCODING THE SMALL SUBUNIT OF RIBULOSE-  
CC BISPHOSPHATE CARBOXYLASE, CHLOROPHYLL A/B BINDING PROTEIN, THE  
CC PROTOCHLOROPHYLLIDE REDUCTASE, RNA, ETC. IT ALSO CONTROLS THE  
CC EXPRESSION OF ITS OWN GENE(S) IN A NEGATIVE FEEDBACK FASHION (BY  
CC SIMILARITY).  
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -!- PFM: CONTAINS ONE COVALENTLY LINKED TETRAPEPTIDE CHROMOPHORE.  
CC -!- SIMILARITY: BELONGS TO THE PHYTOCHROME FAMILY.  
CC -!- SIMILARITY: CONTAINS 2 PAS (PER-ARNT-SIM) DIMERIZATION DOMAINS.  
CC -!- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.  
CC -!- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U56731; AAB41399.1; -;  
CC InterPro; IPR003594; ATPbind\_ATPase.  
CC InterPro; IPR003018; GAF.  
CC InterPro; IPR004359; HIS\_KIN\_sig.  
CC InterPro; IPR003661; His\_KIN.  
CC InterPro; IPR001610; PAC.  
CC InterPro; IPR000014; PAS domain.  
CC InterPro; IPR001294; Phytochrome.  
CC Pfam; PF00360; Phytochrome; 1.  
CC Pfam; PF00512; Signal; 1.  
CC Pfam; PF00989; PAS; 2.  
CC Pfam; PF01590; GAF; 1.  
CC Pfam; PF02518; HATPase\_c; 1.  
CC PRINTS; PR01033; PHYTOCHROME.  
CC SMART; SM00065; GAF; 1.  
CC SMART; SM00387; HATPase\_c; 1.  
CC SMART; SM00388; HSKA; 1.  
CC SMART; SM00086; PAC; 1.  
CC SMART; SM00091; PAS; 2.  
CC TIGRfam; TIGR00229; sensory\_box; 2.  
CC PROSITE; PS50109; HIS\_KIN; 1.  
CC PROSITE; PS50112; PAS; 2.  
CC PROSITE; PS00245; PHYTOCHROME\_1; FALSE\_NEG.  
CC PROSITE; PS50046; PHYTOCHROME\_2; 1.  
CC KW Transcription regulation; Photoreceptor; Phytochrome; Chromophore;  
CC Repeat; Multigene family.  
FT DOMAIN 618 688 PAS 1.  
FT DOMAIN 748 822 PAS 2.  
FT DOMAIN 902 1122 HISTIDINE KINASE.  
FT BINDING 321 321 CHROMOPHORE (BY SIMILARITY).  
SQ SEQUENCE 1135 AA; 126315 MW; AFCC934B7592DE4D CRC64;  
  
Query Match 57.4%; Score 39; DB 1; Length 1135;  
Best Local Similarity 45.5%; Pred. No. 18;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 CFAWQAMRKV 11  
DB 775 CLEWNAKMKI 785  
  
RESULT 8  
ID IL2A\_BOVIN STANDARD; PRT; 275 AA.  
AC P12342;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Interleukin-2 receptor alpha chain precursor (IL-2 receptor alpha  
DE subunit) (P55) (TAC antigen) (CD25).  
GN IL2RA.

OS Bos taurus (Bovine).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 CC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=8212503; PubMed=2835311;  
 RA Weinberg A.D., Shaw J., Paetkau V., Bleackley R.C., Magnuson N.S.,  
 RA Reeves R., Magnuson J.A.;  
 RT "Cloning of cDNA for the bovine IL-2 receptor (bovine Tac antigen).";  
 RL Immunology 63:603-610(1988).  
 RN [2]  
 RP SEQUENCE OF 1-21 FROM N.A.  
 RX MEDLINE=96115968; PubMed=8563178;  
 RA Yoo J., de Leon F.A., Stone R.T., Beattie C.W.;  
 RT "Cloning and chromosomal assignment of the bovine interleukin-2  
 RL receptor alpha (IL-2R alpha) gene.";  
 RL Mamm. Genome 6:751-753(1995).  
 CC [1]  
 CC [2]  
 CC [3]  
 CC [4]  
 CC [5]  
 CC [6]  
 CC [7]  
 CC [8]  
 CC [9]  
 CC [10]  
 CC [11]  
 CC [12]  
 CC [13]  
 CC [14]  
 CC [15]  
 CC [16]  
 CC [17]  
 CC [18]  
 CC [19]  
 CC [20]  
 CC [21]  
 CC [22]  
 CC [23]  
 CC [24]  
 CC [25]  
 CC [26]  
 CC [27]  
 CC [28]  
 CC [29]  
 CC [30]  
 CC [31]  
 CC [32]  
 CC [33]  
 CC [34]  
 CC [35]  
 CC [36]  
 CC [37]  
 CC [38]  
 CC [39]  
 CC [40]  
 CC [41]  
 CC [42]  
 CC [43]  
 CC [44]  
 CC [45]  
 CC [46]  
 CC [47]  
 CC [48]  
 CC [49]  
 CC [50]  
 CC [51]  
 CC [52]  
 CC [53]  
 CC [54]  
 CC [55]  
 CC [56]  
 CC [57]  
 CC [58]  
 CC [59]  
 CC [60]  
 CC [61]  
 CC [62]  
 CC [63]  
 CC [64]  
 CC [65]  
 CC [66]  
 CC [67]  
 CC [68]  
 CC [69]  
 CC [70]  
 CC [71]  
 CC [72]  
 CC [73]  
 CC [74]  
 CC [75]  
 CC [76]  
 CC [77]  
 CC [78]  
 CC [79]  
 CC [80]  
 CC [81]  
 CC [82]  
 CC [83]  
 CC [84]  
 CC [85]  
 CC [86]  
 CC [87]  
 CC [88]  
 CC [89]  
 CC [90]  
 CC [91]  
 CC [92]  
 CC [93]  
 CC [94]  
 CC [95]  
 CC [96]  
 CC [97]  
 CC [98]  
 CC [99]  
 CC [100]  
 CC [101]  
 CC [102]  
 CC [103]  
 CC [104]  
 CC [105]  
 CC [106]  
 CC [107]  
 CC [108]  
 CC [109]  
 CC [110]  
 CC [111]  
 CC [112]  
 CC [113]  
 CC [114]  
 CC [115]  
 CC [116]  
 CC [117]  
 CC [118]  
 CC [119]  
 CC [120]  
 CC [121]  
 CC [122]  
 CC [123]  
 CC [124]  
 CC [125]  
 CC [126]  
 CC [127]  
 CC [128]  
 CC [129]  
 CC [130]  
 CC [131]  
 CC [132]  
 CC [133]  
 CC [134]  
 CC [135]  
 CC [136]  
 CC [137]  
 CC [138]  
 CC [139]  
 CC [140]  
 CC [141]  
 CC [142]  
 CC [143]  
 CC [144]  
 CC [145]  
 CC [146]  
 CC [147]  
 CC [148]  
 CC [149]  
 CC [150]  
 CC [151]  
 CC [152]  
 CC [153]  
 CC [154]  
 CC [155]  
 CC [156]  
 CC [157]  
 CC [158]  
 CC [159]  
 CC [160]  
 CC [161]  
 CC [162]  
 CC [163]  
 CC [164]  
 CC [165]  
 CC [166]  
 CC [167]  
 CC [168]  
 CC [169]  
 CC [170]  
 CC [171]  
 CC [172]  
 CC [173]  
 CC [174]  
 CC [175]  
 CC [176]  
 CC [177]  
 CC [178]  
 CC [179]  
 CC [180]  
 CC [181]  
 CC [182]  
 CC [183]  
 CC [184]  
 CC [185]  
 CC [186]  
 CC [187]  
 CC [188]  
 CC [189]  
 CC [190]  
 CC [191]  
 CC [192]  
 CC [193]  
 CC [194]  
 CC [195]  
 CC [196]  
 CC [197]  
 CC [198]  
 CC [199]  
 CC [200]  
 CC [201]  
 CC [202]  
 CC [203]  
 CC [204]  
 CC [205]  
 CC [206]  
 CC [207]  
 CC [208]  
 CC [209]  
 CC [210]  
 CC [211]  
 CC [212]  
 CC [213]  
 CC [214]  
 CC [215]  
 CC [216]  
 CC [217]  
 CC [218]  
 CC [219]  
 CC [220]  
 CC [221]  
 CC [222]  
 CC [223]  
 CC [224]  
 CC [225]  
 CC [226]  
 CC [227]  
 CC [228]  
 CC [229]  
 CC [230]  
 CC [231]  
 CC [232]  
 CC [233]  
 CC [234]  
 CC [235]  
 CC [236]  
 CC [237]  
 CC [238]  
 CC [239]  
 CC [240]  
 CC [241]  
 CC [242]  
 CC [243]  
 CC [244]  
 CC [245]  
 CC [246]  
 CC [247]  
 CC [248]  
 CC [249]  
 CC [250]  
 CC [251]  
 CC [252]  
 CC [253]  
 CC [254]  
 CC [255]  
 CC [256]  
 CC [257]  
 CC [258]  
 CC [259]  
 CC [260]  
 CC [261]  
 CC [262]  
 CC [263]  
 CC [264]  
 CC [265]  
 CC [266]  
 CC [267]  
 CC [268]  
 CC [269]  
 CC [270]  
 CC [271]  
 CC [272]  
 CC [273]  
 CC [274]  
 CC [275]  
 CC [276]  
 CC [277]  
 CC [278]  
 CC [279]  
 CC [280]  
 CC [281]  
 CC [282]  
 CC [283]  
 CC [284]  
 CC [285]  
 CC [286]  
 CC [287]  
 CC [288]  
 CC [289]  
 CC [290]  
 CC [291]  
 CC [292]  
 CC [293]  
 CC [294]  
 CC [295]  
 CC [296]  
 CC [297]  
 CC [298]  
 CC [299]  
 CC [300]  
 CC [301]  
 CC [302]  
 CC [303]  
 CC [304]  
 CC [305]  
 CC [306]  
 CC [307]  
 CC [308]  
 CC [309]  
 CC [310]  
 CC [311]  
 CC [312]  
 CC [313]  
 CC [314]  
 CC [315]  
 CC [316]  
 CC [317]  
 CC [318]  
 CC [319]  
 CC [320]  
 CC [321]  
 CC [322]  
 CC [323]  
 CC [324]  
 CC [325]  
 CC [326]  
 CC [327]  
 CC [328]  
 CC [329]  
 CC [330]  
 CC [331]  
 CC [332]  
 CC [333]  
 CC [334]  
 CC [335]  
 CC [336]  
 CC [337]  
 CC [338]  
 CC [339]  
 CC [340]  
 CC [341]  
 CC [342]  
 CC [343]  
 CC [344]  
 CC [345]  
 CC [346]  
 CC [347]  
 CC [348]  
 CC [349]  
 CC [350]  
 CC [351]  
 CC [352]  
 CC [353]  
 CC [354]  
 CC [355]  
 CC [356]  
 CC [357]  
 CC [358]  
 CC [359]  
 CC [360]  
 CC [361]  
 CC [362]  
 CC [363]  
 CC [364]  
 CC [365]  
 CC [366]  
 CC [367]  
 CC [368]  
 CC [369]  
 CC [370]  
 CC [371]  
 CC [372]  
 CC [373]  
 CC [374]  
 CC [375]  
 CC [376]  
 CC [377]  
 CC [378]  
 CC [379]  
 CC [380]  
 CC [381]  
 CC [382]  
 CC [383]  
 CC [384]  
 CC [385]  
 CC [386]  
 CC [387]  
 CC [388]  
 CC [389]  
 CC [390]  
 CC [391]  
 CC [392]  
 CC [393]  
 CC [394]  
 CC [395]  
 CC [396]  
 CC [397]  
 CC [398]  
 CC [399]  
 CC [400]  
 CC [401]  
 CC [402]  
 CC [403]  
 CC [404]  
 CC [405]  
 CC [406]  
 CC [407]  
 CC [408]  
 CC [409]  
 CC [410]  
 CC [411]  
 CC [412]  
 CC [413]  
 CC [414]  
 CC [415]  
 CC [416]  
 CC [417]  
 CC [418]  
 CC [419]  
 CC [420]  
 CC [421]  
 CC [422]  
 CC [423]  
 CC [424]  
 CC [425]  
 CC [426]  
 CC [427]  
 CC [428]  
 CC [429]  
 CC [430]  
 CC [431]  
 CC [432]  
 CC [433]  
 CC [434]  
 CC [435]  
 CC [436]  
 CC [437]  
 CC [438]  
 CC [439]  
 CC [440]  
 CC [441]  
 CC [442]  
 CC [443]  
 CC [444]  
 CC [445]  
 CC [446]  
 CC [447]  
 CC [448]  
 CC [449]  
 CC [450]  
 CC [451]  
 CC [452]  
 CC [453]  
 CC [454]  
 CC [455]  
 CC [456]  
 CC [457]  
 CC [458]  
 CC [459]  
 CC [460]  
 CC [461]  
 CC [462]  
 CC [463]  
 CC [464]  
 CC [465]  
 CC [466]  
 CC [467]  
 CC [468]  
 CC [469]  
 CC [470]  
 CC [471]  
 CC [472]  
 CC [473]  
 CC [474]  
 CC [475]  
 CC [476]  
 CC [477]  
 CC [478]  
 CC [479]  
 CC [480]  
 CC [481]  
 CC [482]  
 CC [483]  
 CC [484]  
 CC [485]  
 CC [486]  
 CC [487]  
 CC [488]  
 CC [489]  
 CC [490]  
 CC [491]  
 CC [492]  
 CC [493]  
 CC [494]  
 CC [495]  
 CC [496]  
 CC [497]  
 CC [498]  
 CC [499]  
 CC [500]  
 CC [501]  
 CC [502]  
 CC [503]  
 CC [504]  
 CC [505]  
 CC [506]  
 CC [507]  
 CC [508]  
 CC [509]  
 CC [510]  
 CC [511]  
 CC [512]  
 CC [513]  
 CC [514]  
 CC [515]  
 CC [516]  
 CC [517]  
 CC [518]  
 CC [519]  
 CC [520]  
 CC [521]  
 CC [522]  
 CC [523]  
 CC [524]  
 CC [525]  
 CC [526]  
 CC [527]  
 CC [528]  
 CC [529]  
 CC [530]  
 CC [531]  
 CC [532]  
 CC [533]  
 CC [534]  
 CC [535]  
 CC [536]  
 CC [537]  
 CC [538]  
 CC [539]  
 CC [540]  
 CC [541]  
 CC [542]  
 CC [543]  
 CC [544]  
 CC [545]  
 CC [546]  
 CC [547]  
 CC [548]  
 CC [549]  
 CC [550]  
 CC [551]  
 CC [552]  
 CC [553]  
 CC [554]  
 CC [555]  
 CC [556]  
 CC [557]  
 CC [558]  
 CC [559]  
 CC [560]  
 CC [561]  
 CC [562]  
 CC [563]  
 CC [564]  
 CC [565]  
 CC [566]  
 CC [567]  
 CC [568]  
 CC [569]  
 CC [570]  
 CC [571]  
 CC [572]  
 CC [573]  
 CC [574]  
 CC [575]  
 CC [576]  
 CC [577]  
 CC [578]  
 CC [579]  
 CC [580]  
 CC [581]  
 CC [582]  
 CC [583]  
 CC [584]  
 CC [585]  
 CC [586]  
 CC [587]  
 CC [588]  
 CC [589]  
 CC [590]  
 CC [591]  
 CC [592]  
 CC [593]  
 CC [594]  
 CC [595]  
 CC [596]  
 CC [597]  
 CC [598]  
 CC [599]  
 CC [600]  
 CC [601]  
 CC [602]  
 CC [603]  
 CC [604]  
 CC [605]  
 CC [606]  
 CC [607]  
 CC [608]  
 CC [609]  
 CC [610]  
 CC [611]  
 CC [612]  
 CC [613]  
 CC [614]  
 CC [615]  
 CC [616]  
 CC [617]  
 CC [618]  
 CC [619]  
 CC [620]  
 CC [621]  
 CC [622]  
 CC [623]  
 CC [624]  
 CC [625]  
 CC [626]  
 CC [627]  
 CC [628]  
 CC [629]  
 CC [630]  
 CC [631]  
 CC [632]  
 CC [633]  
 CC [634]  
 CC [635]  
 CC [636]  
 CC [637]  
 CC [638]  
 CC [639]  
 CC [640]  
 CC [641]  
 CC [642]  
 CC [643]  
 CC [644]  
 CC [645]  
 CC [646]  
 CC [647]  
 CC [648]  
 CC [649]  
 CC [650]  
 CC [651]  
 CC [652]  
 CC [653]  
 CC [654]  
 CC [655]  
 CC [656]  
 CC [657]  
 CC [658]  
 CC [659]  
 CC [660]  
 CC [661]  
 CC [662]  
 CC [663]  
 CC [664]  
 CC [665]  
 CC [666]  
 CC [667]  
 CC [668]  
 CC [669]  
 CC [670]  
 CC [671]  
 CC [672]  
 CC [673]  
 CC [674]  
 CC [675]  
 CC [676]  
 CC [677]  
 CC [678]  
 CC [679]  
 CC [680]  
 CC [681]  
 CC [682]  
 CC [683]  
 CC [684]  
 CC [685]  
 CC [686]  
 CC [687]  
 CC [688]  
 CC [689]  
 CC [690]  
 CC [691]  
 CC [692]  
 CC [693]  
 CC [694]  
 CC [695]  
 CC [696]  
 CC [697]  
 CC [698]  
 CC [699]  
 CC [700]  
 CC [701]  
 CC [702]  
 CC [703]  
 CC [704]  
 CC [705]  
 CC [706]  
 CC [707]  
 CC [708]  
 CC [709]  
 CC [710]  
 CC [711]  
 CC [712]  
 CC [713]  
 CC [714]  
 CC [715]  
 CC [716]  
 CC [717]  
 CC [718]  
 CC [719]  
 CC [720]  
 CC [721]  
 CC [722]  
 CC [723]  
 CC [724]  
 CC [725]  
 CC [726]  
 CC [727]  
 CC [728]  
 CC [729]  
 CC [730]  
 CC [731]  
 CC [732]  
 CC [733]  
 CC [734]  
 CC [735]  
 CC [736]  
 CC [737]  
 CC [738]  
 CC [739]  
 CC [740]  
 CC [741]  
 CC [742]  
 CC [743]  
 CC [744]  
 CC [745]  
 CC [746]  
 CC [747]  
 CC [748]  
 CC [749]  
 CC [750]  
 CC [751]  
 CC [752]  
 CC [753]  
 CC [754]  
 CC [755]  
 CC [756]  
 CC [757]  
 CC [758]  
 CC [759]  
 CC [760]  
 CC [761]  
 CC [762]  
 CC [763]  
 CC [764]  
 CC [765]  
 CC [766]  
 CC [767]  
 CC [768]  
 CC [769]  
 CC [770]  
 CC [771]  
 CC [772]  
 CC [773]  
 CC [774]  
 CC [775]  
 CC [776]  
 CC [777]  
 CC [778]  
 CC [779]  
 CC [780]  
 CC [781]  
 CC [782]  
 CC [783]  
 CC [784]  
 CC [785]  
 CC [786]  
 CC [787]  
 CC [788]  
 CC [789]  
 CC [790]  
 CC [791]  
 CC [792]  
 CC [793]  
 CC [794]  
 CC [795]  
 CC [796]  
 CC [797]  
 CC [798]  
 CC [799]  
 CC [800]  
 CC [801]  
 CC [802]  
 CC [803]  
 CC [804]  
 CC [805]  
 CC [806]  
 CC [807]  
 CC [808]  
 CC [809]  
 CC [810]  
 CC [811]  
 CC [812]  
 CC [813]  
 CC [814]  
 CC [815]  
 CC [816]  
 CC [817]  
 CC [818]  
 CC [819]  
 CC [820]  
 CC [821]  
 CC [822]  
 CC [823]  
 CC [824]  
 CC [825]  
 CC [826]  
 CC [827]  
 CC [828]  
 CC [829]  
 CC [830]  
 CC [831]  
 CC [832]  
 CC [833]  
 CC [834]  
 CC [835]  
 CC [836]  
 CC [837]  
 CC [838]  
 CC [839]  
 CC [840]  
 CC [841]  
 CC [842]  
 CC [843]  
 CC [844]  
 CC [845]  
 CC [846]  
 CC [847]  
 CC [848]  
 CC [849]  
 CC [850]  
 CC [851]  
 CC [852]  
 CC [853]  
 CC [854]  
 CC [855]  
 CC [856]  
 CC [857]  
 CC [858]  
 CC [859]  
 CC [860]  
 CC [861]  
 CC [862]  
 CC [863]  
 CC [864]  
 CC [865]  
 CC [866]  
 CC [867]  
 CC [868]  
 CC [869]  
 CC [870]  
 CC [871]  
 CC [872]  
 CC [873]  
 CC [874]  
 CC [875]  
 CC [876]  
 CC [877]  
 CC [878]  
 CC [879]  
 CC [880]  
 CC [881]  
 CC [882]  
 CC [883]  
 CC [884]  
 CC [885]  
 CC [886]  
 CC [887]  
 CC [888]  
 CC [889]  
 CC [890]  
 CC [891]  
 CC [892]  
 CC [893]  
 CC [894]  
 CC [895]  
 CC [896]  
 CC [897]  
 CC [898]  
 CC [899]  
 CC [900]  
 CC [901]  
 CC [902]  
 CC [903]  
 CC [904]  
 CC [905]  
 CC [906]  
 CC [907]  
 CC [908]  
 CC [909]  
 CC [910]  
 CC [911]  
 CC [912]  
 CC [913]  
 CC [914]  
 CC [915]  
 CC [916]  
 CC [917]  
 CC [918]  
 CC [919]  
 CC [920]  
 CC [921]  
 CC [922]  
 CC [923]  
 CC [924]  
 CC [925]  
 CC [926]  
 CC [927]  
 CC [928]  
 CC [929]  
 CC [930]  
 CC [931]  
 CC [932]  
 CC [933]  
 CC [934]  
 CC [935]  
 CC [936]  
 CC [937]  
 CC [938]  
 CC [939]  
 CC [940]  
 CC [941]  
 CC [942]  
 CC [943]  
 CC [944]  
 CC [945]  
 CC [946]  
 CC [947]  
 CC [948]  
 CC [949]  
 CC [950]  
 CC [951]  
 CC [952]  
 CC [953]  
 CC [954]  
 CC [955]  
 CC [956]  
 CC [957]  
 CC [958]  
 CC [959]  
 CC [960]  
 CC [961]  
 CC [962]  
 CC [963]  
 CC [964]  
 CC [965]  
 CC [966]  
 CC [967]  
 CC [968]  
 CC [969]  
 CC [970]  
 CC [971]  
 CC [972]  
 CC [973]  
 CC [974]  
 CC [975]  
 CC [976]  
 CC [977]  
 CC [978]  
 CC [979]  
 CC [980]  
 CC [981]  
 CC [982]  
 CC [983]  
 CC [984]  
 CC [985]  
 CC [986]  
 CC [987]  
 CC [988]  
 CC [989]  
 CC [990]  
 CC [991]  
 CC [992]  
 CC [993]  
 CC [994]  
 CC [995]  
 CC [996]  
 CC [997]  
 CC [998]  
 CC [999]  
 CC [1000]  
 CC [1001]  
 CC [1002]  
 CC [1003]  
 CC [1004]  
 CC [1005]  
 CC [1006]  
 CC [1007]  
 CC [1008]  
 CC [1009]  
 CC [1010]  
 CC [1011]  
 CC [1012]  
 CC [1013]  
 CC [1014]  
 CC [1015]  
 CC [1016]  
 CC [1017]  
 CC [1018]  
 CC [1019]  
 CC [1020]  
 CC [1021]  
 CC [1022]  
 CC [1023]  
 CC [1024]  
 CC [1025]  
 CC [1026]  
 CC [1027]  
 CC [1028]  
 CC [1029]  
 CC [1030]  
 CC [1031]  
 CC [1032]  
 CC [1033]  
 CC [1034]  
 CC [1035]  
 CC [1036]  
 CC [1037]  
 CC [1038]  
 CC [1039]  
 CC [1040]  
 CC [1041]  
 CC [1042]  
 CC [1043]  
 CC [1044]  
 CC [1045]  
 CC [1046]  
 CC [1047]  
 CC [1048]  
 CC [1049]  
 CC [1050]  
 CC [1051]  
 CC [1052]  
 CC [1053]  
 CC [1054]  
 CC [1055]  
 CC [1056]  
 CC [1057]  
 CC [1058]  
 CC [1059]  
 CC [1060]  
 CC [1061]  
 CC [1062]  
 CC [1063]  
 CC [1064]  
 CC [1065]  
 CC [1066]  
 CC [1067]  
 CC [1068]  
 CC [1069]  
 CC [1070]  
 CC [1071]  
 CC [1072]  
 CC [1073]  
 CC [1074]  
 CC [1075]  
 CC [1076]  
 CC [1077]  
 CC [1078]  
 CC [1079]  
 CC [1080]  
 CC [1081]  
 CC [1082]  
 CC [1083]  
 CC [1084]  
 CC [1085]  
 CC [1086]  
 CC [1087]  
 CC [1088]  
 CC [1089]  
 CC [1090]  
 CC [1091]  
 CC [1092]  
 CC [1093]  
 CC [1094]  
 CC [1095]  
 CC [1096]  
 CC [1097]  
 CC [1098]  
 CC [1099]  
 CC [1100]  
 CC [1101]  
 CC [1102]  
 CC [1103]  
 CC [1104]  
 CC [1105]  
 CC [1106]  
 CC [1107]  
 CC [1108]  
 CC [1109]  
 CC [1110]  
 CC [1111]  
 CC [1112]  
 CC [1113]  
 CC [1114]  
 CC [1115]  
 CC [1116]  
 CC [1117]  
 CC [1118]  
 CC [1119]  
 CC [1120]  
 CC [1121]  
 CC [1122]  
 CC [1123]  
 CC [1124]  
 CC [1125]  
 CC [1126]  
 CC [1127]  
 CC [1128]  
 CC [1129]  
 CC [1130]  
 CC [1131]  
 CC [1132]  
 CC [1133]  
 CC [1134]  
 CC [1135]  
 CC [1136]  
 CC [1137]  
 CC [1138]  
 CC [1139]  
 CC [1140]  
 CC [1141]  
 CC [1142]  
 CC [1143]  
 CC [1144]  
 CC [1145]  
 CC [1146]  
 CC [1147]  
 CC [1148]  
 CC [1149]  
 CC [1150]  
 CC [1151]  
 CC [1152]  
 CC [1153]  
 CC [1154]  
 CC [1155]  
 CC [1156]  
 CC [1157]  
 CC [1158]  
 CC [1159]  
 CC [1160]  
 CC [1161]  
 CC [1162]  
 CC [1163]  
 CC [1164]  
 CC [1165]  
 CC [1166]  
 CC [1167]  
 CC [1168]  
 CC [1169]  
 CC [1170]  
 CC [1171]  
 CC [1172]  
 CC [1173]  
 CC [1174]  
 CC [1175]  
 CC [1176]  
 CC [1177]



```

RESULT 10
ID UL24_HSVSA STANDARD; PRT; 303 AA.
AC Q01005;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Gene 20 protein.
GN 20.
OS Herpesvirus saimiri (strain 11).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
CX NCBI_TaxID=10383;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92333688; PubMed=1321287;
RA Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,
RA Newman C., Wittmann S., Craxton M.A., Coleman H., Fleckenstein B.,
RA Honess R.W.;
RT "Primary structure of the herpesvirus saimiri genome.";
RL J. Virol. 66:5047-5058(1992).
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL24,
CC - EHV-1 37, EBV BXRFL, HCMV UL76, ILTV ORF3, AND VZV 35.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X64346; CAA45644.1; -
CC InterPro; IPR002580; WZBEM6.
CC PIR; B36808; WZBEM6.
CC InterPro; IPR002580; Herpes UL24.
CC Pfam; PF01646; Herpes UL24; 1.
CC SEQUENCE 303 AA; 34942 MW; DF6D59F7A1C93A0B CRC64;
CC
CC Query Match 54.4%; Score 37; DB 1; Length 303;
CC Best Local Similarity 50.0%; Pred. No. 11;
CC Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
CC
CC QY 1 CFWRQAMRKVR 12
CC ||| ||| |||
CC Db 174 CFLMSRADVEIR 185
CC
CC RESULT 11
ID WD66_PHYPO STANDARD; PRT; 601 AA.
AC P90587;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 66 kDa stress protein (p66).
OS Physarum polycephalum (Slime mold).
OC Eukaryota; Mycetozoa; Myxogastria; Myxogastromycetidae; Physarida;
OC Physarum.
CX NCBI_TaxID=5791;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98351989; PubMed=9685722;
RA Matsumoto S., Ogawa M., Kasakura T., Shimada Y., Mitsui M., Maruya M.,
RA Isohata M., Yahara I., Murakami-Murofushi K.;
RT "A novel 66-kDa stress protein, p66, associated with the process of
RT cyst formation of Physarum polycephalum is a Physarum homologue of a
RT yeast actin-interacting protein, AIP1.";
RL J. Biochem. 124:326-331(1998).
CC -!- FUNCTION: ASSOCIATED WITH THE PROCESS OF CYST FORMATION.
CC -!- SIMILARITY: CONTAINS 10 WD REPEATS (TRP-ASP DOMAINS).
CC -!- SIMILARITY: BELONGS TO THE AIP1 FAMILY OF WD-REPEAT PROTEINS.
CC

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U86011; AAC26321.1; -
CC InterPro; IPR001880; WD40.
CC Pfam; PF00400; WD40; 10.
CC SMART; PRO0320; GPROTEINBRPT.
CC SMART; SM00320; WD40; 10.
CC PROSITE; PS00678; WD_REPEATS_1; FALSE_NEG.
CC PROSITE; PS50082; WD_REPEATS_2; 6.
CC PROSITE; PS50294; WD_REPEATS_REGION; 1.
CC Repeat; WD repeat.
KW REPEAT 56 95 WD 1.
FT REPEAT 100 143 WD 2.
FT REPEAT 145 184 WD 3.
FT REPEAT 187 226 WD 4.
FT REPEAT 233 272 WD 5.
FT REPEAT 318 357 WD 6.
FT REPEAT 435 478 WD 7.
FT REPEAT 493 522 WD 8.
FT REPEAT 526 565 WD 9.
FT REPEAT 569 600 WD 10.
SQ SEQUENCE 601 AA; 64321 MW; F691217D838F747A CRC64;
CC
CC Query Match 54.4%; Score 37; DB 1; Length 601;
CC Best Local Similarity 45.5%; Pred. No. 23;
CC Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
CC
CC QY 2 FAWQRAMRKVR 12
CC ||| ||| |||
CC Db 510 FVWDKASRKIK 520
CC
CC RESULT 12
ID PHYC_ORYSA STANDARD; PRT; 1137 AA.
AC Q9ZWI9; P93429;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phytochrome C.
GN PHYC.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
CX NCBI_TaxID=45330;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Tahir M., Kanegae H., Takano M.;
RT "Phytochrome C (PHYC) gene in rice: isolation and characterization of
RT a complete coding sequence.";
RL (in) Plant Gene Register PGR98-210.
RN [2]
RP SEQUENCE OF 275-378 FROM N.A.
RX MEDLINE=97019052; PubMed=8865668;
RA Mathews S., Sharrock R.A.;
RT "The phytochrome gene family in grasses (Poaceae): a phylogeny and
RT evidence that grasses have a subset of the loci found in dicot
RT angiosperms.";
RL Mol. Biol. Evol. 13:1141-1150(1996).
CC -!- FUNCTION: REGULATORY PHOTORECEPTOR WHICH EXISTS IN TWO FORMS THAT
CC ARE REVERSIBLY INTERCONVERTIBLE BY LIGHT: THE PR FORM THAT ABSORBS
CC MAXIMALLY IN THE RED REGION OF THE SPECTRUM AND THE PFR FORM THAT
CC ABSORBS MAXIMALLY IN THE FAR-RED REGION. PHOTOCONVERSION OF PR IN
CC PFR INDUCES AN ARRAY OF MORPHOGENIC RESPONSES, WHEREAS
CC RECONVERSION OF PFR TO PR CANCELS THE INDUCTION OF THOSE

```

CC RESPONSES. PER CONTROLS THE EXPRESSION OF A NUMBER OF NUCLEAR  
 CC GENES INCLUDING THOSE ENCODING THE SMALL SUBUNIT OF RIBULOSE-  
 CC BIPHOSPHATE CARBOXYLASE, CHLOROPHYLL A/B BINDING PROTEIN,  
 CC PROTOCHLOROPHYLLIDE REDUCTASE, RNA, ETC. IT ALSO CONTROLS  
 CC THE EXPRESSION OF ITS OWN GENE(S) IN A NEGATIVE FEEDBACK FASHION.  
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -1- P-TM: CONTAINS ONE COVALENTLY LINKED TETRAPEPTIDE CHROMOPHORE.  
 CC -1- SIMILARITY: BELONGS TO THE PHYTOCHROME FAMILY.  
 CC -1- SIMILARITY: CONTAINS 2 PAS (PER-ARNT-SIM) DIMERIZATION DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 CC EMBL; AB018442; BAA74448.1; -  
 CC EMBL; U61207; ABA41996.1; -  
 CC InterPro; IPR003594; ATPbind\_ATPase.  
 CC InterPro; IPR003018; GAR.  
 CC InterPro; IPR004359; HIS\_KIN\_sig.  
 CC InterPro; IPR003661; HIS\_KIN\_A.  
 CC InterPro; IPR001610; PAC.  
 CC InterPro; IPR000014; PAS domain.  
 CC InterPro; IPR001294; Phytochrome.  
 CC Pfam; PF00360; phytochrome; 1.  
 CC Pfam; PF00512; signal; 1.  
 CC Pfam; PF00989; PAS; 2.  
 CC Pfam; PF01590; GAR; 1.  
 CC Pfam; PF02518; HATPase\_C; 1.  
 CC PRINTS; PR01033; PHYTOCHROME.  
 CC SMART; SM00065; GAR; 1.  
 CC SMART; SM00387; HATPase\_C; 1.  
 CC SMART; SM00388; HSKA; 1.  
 CC SMART; SM00086; PAC; 1.  
 CC SMART; SM00091; PAS; 2.  
 CC TIGRFAMs; TIGR00229; sensory\_box; 2.  
 CC PROSITE; PS50109; HIS\_KIN; 1.  
 CC PROSITE; PS50112; PAS; 2.  
 CC PROSITE; PS00245; PHYTOCHROME\_1; 1.  
 CC PROSITE; PS50046; PHYTOCHROME\_2; 1.  
 CC Repeat; Multigene family.  
 CC Domain; 620 690 PAS 1.  
 CC Domain; 750 824 PAS 2.  
 CC Domain; 904 1124 HISTIDINE KINASE.  
 CC BINDING 322 322 CHROMOPHORE (BY SIMILARITY).  
 CC CONFLICT 279 279 F -> S (IN REF. 2).  
 CC CONFLICT 292 292 C -> S (IN REF. 2).  
 CC SEQUENCE 1137 AA; 125982 MW; F2A520181CFE7B32 CRC64;  
 CC  
 CC Query Match 54.4%; Score 37; DB 1; Length 1137;  
 CC Best Local Similarity 45.5%; Pred. No. 44;  
 CC Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 CC  
 CC QY 1 CFAWQANRKY 11  
 CC DB 777 CLENNANMOKI 787  
 CC  
 CC RESULT 13  
 CC ATX1\_ARATH STANDARD; PRT; 1179 AA.  
 CC ID ATX1\_ARATH  
 CC AC Q9LT02;  
 CC DT 16-OCT-2001 (Rel. 40, Created)  
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 CC DE Potential cation-transporting ATPase (EC 3.6.3.-).  
 CC GN AT5G23630 OR MQM1.11.

OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eursids II; Brassicales; Brassicaceae; Arabidopsi.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=20181125; PubMed=10718197;  
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,  
 RA Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence  
 RT features of the regions of 3,076,755 bp covered by sixty PI and TAC  
 RL clones.";  
 RL DNA Res. 7:31-63(2000).  
 CC -1- FUNCTION: INVOLVED IN TRANSPORT OF CATIONS (POTENTIAL).  
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY  
 CC (E1-E2 ATPASES). SUBFAMILY V.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 CC EMBL; AB025633; BAA97238.1; -  
 CC InterPro; IPR001757; ATPase\_E1-E2.  
 CC InterPro; IPR001454; H1gnase/hydrlase.  
 CC Pfam; PF00122; E1-E2\_ATPase; 1.  
 CC Pfam; PF00702; Hydrolase; 1.  
 CC PRINTS; PR00119; CAIATPASE.  
 CC PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
 CC Hydrolase; Transmembrane; Phosphorylation; ATP-binding; Metal-binding;  
 CC Magnesium.  
 CC Domain; 1 20 CYTOPLASMIC (POTENTIAL).  
 CC TRANSMEM 21 42 POTENTIAL.  
 CC Domain; 43 50 EXTRACELLULAR (POTENTIAL).  
 CC TRANSMEM 51 71 POTENTIAL.  
 CC Domain; 72 192 CYTOPLASMIC (POTENTIAL).  
 CC TRANSMEM 193 215 POTENTIAL.  
 CC Domain; 216 218 EXTRACELLULAR (POTENTIAL).  
 CC TRANSMEM 219 238 POTENTIAL.  
 CC Domain; 239 402 CYTOPLASMIC (POTENTIAL).  
 CC TRANSMEM 403 422 POTENTIAL.  
 CC Domain; 423 435 EXTRACELLULAR (POTENTIAL).  
 CC TRANSMEM 436 453 POTENTIAL.  
 CC Domain; 454 947 CYTOPLASMIC (POTENTIAL).  
 CC TRANSMEM 948 967 POTENTIAL.  
 CC Domain; 968 979 EXTRACELLULAR (POTENTIAL).  
 CC TRANSMEM 980 997 POTENTIAL.  
 CC Domain; 998 1013 CYTOPLASMIC (POTENTIAL).  
 CC TRANSMEM 1014 1034 POTENTIAL.  
 CC Domain; 1035 1059 EXTRACELLULAR (POTENTIAL).  
 CC TRANSMEM 1060 1079 POTENTIAL.  
 CC Domain; 1080 1092 CYTOPLASMIC (POTENTIAL).  
 CC TRANSMEM 1093 1110 POTENTIAL.  
 CC Domain; 1111 1128 EXTRACELLULAR (POTENTIAL).  
 CC TRANSMEM 1129 1148 POTENTIAL.  
 CC Domain; 1149 1179 POTENTIAL.  
 CC TRANSMEM 1175 1178 CYTOPLASMIC (POTENTIAL).  
 CC Domain; 812 812 POLY-LYS.  
 CC TRANSMEM 816 816 PHOSPHORYLATION (BY SIMILARITY).  
 CC METAL 812 812 MAGNESIUM (BY SIMILARITY).  
 CC METAL 816 816 MAGNESIUM (BY SIMILARITY).  
 CC SEQUENCE 1179 AA; 131115 MW; 4A3E82D222A4D78 CRC64;  
 CC  
 CC Query Match 54.4%; Score 37; DB 1; Length 1179;  
 CC Best Local Similarity 44.4%; Pred. No. 45;  
 CC Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFAWQAMR 9  
DB 1144 CYSWERLRL 1152

RESULT 14  
ABCR HUMAN  
ID ABCR HUMAN STANDARD, PRT; 2273 AA.  
AC P78363; 060438; 060915; 015112;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DE 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Retinal-specific ATP-binding cassette transporter (RIM ABC  
DE transporter) (RIM protein) (RMP) (Stargardt disease protein).  
GN ABCA4 OR ABCR.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RN SEQUENCE FROM N.A., VARIANTS STGD, AND VARIANTS HIS-846 AND GLN-943.  
RP MEDLINE=97207541; PubMed=9054934;  
RA Allikmets R., Singh N., Sun H., Shroyer N.F., Hutchinson A.,  
RA Chidambaram A., Gerard B., Baird L., Stauffer D., Peiffer A.,  
RA Rattner A., Smallwood P.M., Li Y., Anderson K.L., Lewis R.A.,  
RA Nathans J., Leppert M., Dean M., Lupski J.R.;  
RA "A photoreceptor cell-specific ATP-binding transporter gene (ABCR) is  
RT mutated in recessive Stargardt macular dystrophy.";  
RL Nat. Genet. 15:236-246 (1997).  
RN [2]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=97345663; PubMed=9202155;  
RA Azarian S.M., Travis G.H.;  
RA "The photoreceptor rim protein is an ABC transporter encoded by the  
RT gene for recessive Stargardt's disease (ABCR).";  
RL FEBS Lett. 409:247-252 (1997).  
RN [3]  
RN SEQUENCE FROM N.A., AND VARIANTS STGD TRP-18 AND CYS-212.  
RP MEDLINE=98163759; PubMed=9503029;  
RA Glaner S., Rozet J.-M., van de Pol T.J.R., Hoyng C.B., Munnich A.,  
RA Blankenagel A., Kaplan J., Cremers F.P.M.;  
RA "Complete exon-intron structure of the retina-specific ATP binding  
RT transporter gene (ABCR) allows the identification of novel mutations  
RT underlying Stargardt disease.";  
RL Genomics 48:139-142 (1998).  
RN [4]  
RN SEQUENCE FROM N.A., AND VARIANTS STGD.  
RP MEDLINE=98141123; PubMed=9490294;  
RA Nasonkin I., Illing M., Koehler M.R., Schmid M., Molday R.S.,  
RA Weber B.H.F.;  
RA "Mapping of the rod photoreceptor ABC transporter (ABCR) to 1p21-p22.1  
RT and identification of novel mutations in Stargardt's disease.";  
RL Hum. Genet. 102:21-26 (1998).  
RN [5]  
RN CHARACTERIZATION.  
RP MEDLINE=99175213; PubMed=10075733;  
RA Sun H., Molday R.S., Nathans J.;  
RA "Retinal stimulates ATP hydrolysis by purified and reconstituted ABCR,  
RT the photoreceptor-specific ATP-binding cassette transporter  
RT responsible for Stargardt disease.";  
RL J. Biol. Chem. 274:8269-8281 (1999).  
RN [6]  
RN DISEASE.  
RP MEDLINE=98133912; PubMed=9466990;  
RA Cremers F.P.M., van de Pol D.J.R., van Driel M.A., den Hollander A.I.,  
RA van Haren F.J.J., Knoers N.V.A.M., Tijmes N., Bergen A.A.B.,  
RA Rohrschneider K., Blankenagel A., Pinckers A.J.L.G., Deutman A.F.,  
RA Hoyng C.B.;  
RA "Autosomal recessive retinitis pigmentosa and cone-rod dystrophy  
RT caused by splice site mutations in the Stargardt's disease gene  
RT ABCR.";  
RL Hum. Mol. Genet. 7:355-362 (1998).  
RN [7]

RP VARIANTS ARMD2, AND VARIANTS.  
RX MEDLINE=9742530; PubMed=9295268;  
RA Allikmets R., Shroyer N.F., Singh N., Seddon J.M., Lewis R.A.,  
RA Bernstein P.S., Peiffer A., Zabriskie N.A., Li Y., Hutchinson A.,  
RA Dean M., Lupski J.R., Leppert M.;  
RA "Mutation of the Stargardt disease gene (ABCR) in age-related macular  
RT degeneration.";  
RL Science 277:1805-1807 (1997).  
RN [8]  
RN VARIANTS STGD W-18; C-212; H-636; M-1019; V-1038; C-1108; W-1640;  
RP S-1977 AND H-2107, AND VARIANTS FFM P-11; P-341; V-1038; E-1091;  
RP C-1508; F-1970 AND R-1971.  
RX MEDLINE=98454319; PubMed=9781034;  
RA Rozet J.-M., Gerber S., Souied E., Perrault I., Chatelin S., Ghazi I.,  
RA Leowski C., Dufier J.-L., Munnich A., Kaplan J.;  
RA "Spectrum of ABCR gene mutations in autosomal recessive macular  
RT dystrophies.";  
RL Eur. J. Hum. Genet. 6:291-295 (1998).  
RN [9]  
RN VARIANTS STGD.  
RP MEDLINE=99138655; PubMed=9973280;  
RA Lewis R.A., Shroyer N.F., Singh N., Allikmets R., Hutchinson A.,  
RA Li Y., Lupski J.R., Leppert M., Dean M.;  
RA "Genotype/phenotype analysis of a photoreceptor-specific ATP-binding  
RT cassette transporter gene, ABCR, in Stargardt disease.";  
RL Am. J. Hum. Genet. 64:422-434 (1999).  
RN [10]  
RN VARIANTS STGD, AND VARIANTS.  
RX MEDLINE=99192348; PubMed=10090887;  
RA Maugeri A., van Driel M.A., van de Pol D.J.R., Klevering B.J.,  
RA van Haren F.J.J., Tijmes N., Bergen A.A.B., Rohrschneider K.,  
RA Blankenagel A., Pinckers A.J.L.G., Dahl N., Brunner H.G.,  
RA Deutman A.F., Hoyng C.B., Cremers F.P.M.;  
RA "The 2588G->C mutation in the ABCR gene is a mild frequent founder  
RT mutation in the western European population and allows the  
RT classification of ABCR mutations in patients with Stargardt disease.";  
RL Am. J. Hum. Genet. 64:1024-1035 (1999).  
RN [11]  
RN VARIANT STGD TYR-54, AND VARIANT ALA-863.  
RX MEDLINE=20077755; PubMed=10612508;  
RA Zhang K., Garibaldi D.C., Khazava M., Albini T., Chiang M.F.,  
RA Kerrigan M., Sunness J.S., Han M., Allikmets R.;  
RA "A novel mutation in the ABCR gene in four patients with autosomal  
RT recessive Stargardt disease.";  
RL Am. J. Ophthalmol. 128:720-724 (1999).  
RN [12]  
RN VARIANTS STGD V-60; R-206; N-300; P-541; A-849; P-974; V-1038; C-1108;  
RP L-1408; R-1488; D-1652; P-1729; E-1961; W-2038; W-2077; H-2107; R-2128  
RP AND Y-2150.  
RX MEDLINE=99221420; PubMed=10206579;  
RA Fiehman G.A., Stone E.M., Grover S., Derlacki D.J., Haines H.L.,  
RA Hockey R.R.;  
RA "Variation of clinical expression in patients with Stargardt dystrophy  
RT and sequence variations in the ABCR gene.";  
RL Arch. Ophthalmol. 117:504-510 (1999).  
RN [13]  
RN VARIANTS GLU-1961 AND ASN-2177.  
RX MEDLINE=20349288; PubMed=10830298;  
RA Allikmets R., Tammur J., Hutchinson A., Lewis R.A., Shroyer N.F.,  
RA Datakhvili K., Lupski J.R., Steiner K., Pauleikhoff D., Holz F.G.,  
RA Weber B.H.F., Dean M., Atkinson A., Gail M.H., Bernstein P.S.,  
RA Singh N., Peiffer A., Zabriskie N.A., Leppert M., Seddon J.M.,  
RA Zhang K., Sunness J.S., Udar N.S., Velchits S., Silva-Garcia R.,  
RA Small K.W., Simonelli F., Testa F., D'Urso M., Brancato R.,  
RA Rinaldi E., Ingavst S., Taube A., Wadelius C., Souied E., Ducrocq D.,  
RA Kaplan J., Assink J.J.M., ten Brink J.B., de Jong P.T.V.M.,  
RA Bergen A.A.B., Maugeri A., van Driel M.A., Hoyng C.B., Cremers F.P.M.,  
RA Paloma E., Coco R., Balcells S., Gonzalez-Duarte R., Kermani S.,  
RA Stanga P., Bhatnagary S.S., Bird A.C.;  
RA "Further evidence for an association of ABCR alleles with age-related  
RT macular degeneration.";  
RL Am. J. Hum. Genet. 67:487-491 (2000).  
RN [14]

RP VARIANTS STGD E-60; T-60; E-65; L-68; R-72; C-212; S-230; S-247;  
 RP V-328; K-471; P-541; Q-572; R-607; K-635; C-653; Y-764; R-765; A-901;  
 RP I-959; K-1036; V-1038; P-1063; D-1087; C-1108; L-1380; K-1399;  
 RP P-1430; V-1440; L-1486; Y-1488; M-1537; P-1689; L-1705;  
 RP T-1733; R-1748; P-1763; K-1885; H-1898; E-1961; R-1975; S-1977; G-2077  
 RP W-2077 AND V-2241, AND VARIANTS Q-152; H-212; R-423; I-552; R-914;  
 RP Q-943; T-1562; I-1868; M-1921; L-1948; R-1970; A-2059; N-2177 AND  
 RP V-2216.  
 RX MEDLINE=20442027; PubMed=1058763;  
 RA Rivera A., White K., Stoehr H., Steiner K., Hemmrich N., Grimm T.,  
 RA Jurkies B., Lorenz B., Scholl H.P.N., Apfelstedt-Sylla E.,  
 RA Weber B.H.F.,  
 RT "A comprehensive survey of sequence variation in the ABCA4 (ABCR) gene  
 RT in Stargardt disease and age-related macular degeneration.",  
 RL Am. J. Hum. Genet. 67:1800-1813(2000).  
 RN [15]  
 RP VARIANTS CORD3 GLU-65; CYS-212; PRO-541; ALA-863; GLY-863 DEL;  
 RP VAL-1038; LYS-1122; TYR-1490 AND ASP-1598.  
 RX MEDLINE=20442040; PubMed=10958761;  
 RA Maugeri A., Klovering B., Rohrschneider K., Blankensgel A.,  
 RA Brunner H.G., Deutman A.F., Hoyng C.B., Cremers F.P.M.,  
 RT "Mutations in the ABCA4 (ABCR) gene are the major cause of autosomal  
 RT recessive cone-rod dystrophy.",  
 RL Am. J. Hum. Genet. 67:960-966(2000).  
 RN [16]  
 RP VARIANTS STGD ASP-340; GLN-572; ALA-863; SER-965; VAL-1038; ALA-1780  
 RP AND HIS-1898, AND VARIANT GLN-943.  
 RX MEDLINE=20208356; PubMed=10746567;  
 RA Shroyer N.F., Lewis R.A., Lupski J.R.,  
 RT "Complex inheritance of ABCR mutations in Stargardt disease: linkage  
 RT disequilibrium, complex alleles, and pseudodominance.",  
 RL Hum. Genet. 106:244-248(2000).  
 RN [17]  
 RP VARIANTS STGD.  
 RX MEDLINE=20095082; PubMed=10634594;  
 RA Papaioannou M., Osaka L., Bessant D., Lois N., Bird A.C., Payne A.,  
 RA Bhattacharya S.S.,  
 RT "An analysis of ABCR mutations in British patients with recessive  
 RT retinal dystrophies.",  
 RL Invest. Ophthalmol. Vis. Sci. 41:16-19(2000).  
 RN [18]  
 RP VARIANTS STGD C-212; D-767; I-897; V-1038; K-1087; K-1399; Q-1640 AND  
 RP E-1961, AND VARIANT HIS-212.  
 RX MEDLINE=20174852; PubMed=10711710;  
 RA Simonelli F., Testa F., de Crecchio G., Rinaldi E., Hutchinson A.,  
 RA Atkinson A., Dean M., D'Urso M., Allikmets R.,  
 RT "New ABCR mutations and clinical phenotype in Italian patients with  
 RT Stargardt disease.",  
 RL Invest. Ophthalmol. Vis. Sci. 41:892-897(2000).  
 RN [19]  
 RP CHARACTERIZATION OF VARIANTS, AND MUTAGENESIS OF GLY-966; LYS-969;  
 RP GLY-1975 AND LYS-1978.  
 RX MEDLINE=20472331; PubMed=11017087;  
 RA Sun H., Smallwood P.M., Nathans J.,  
 RT "Biochemical defects in ABCR protein variants associated with human  
 RT retinopathies.",  
 RL Nat. Genet. 26:242-246(2000).  
 RN [20]  
 RP VARIANTS STGD ASN-972, AND VARIANTS GLN-943; ILE-1868 AND LEU-1948.  
 RX MEDLINE=21478761; PubMed=11594993;  
 RA Eksandh L., Ekstrom U., Abrahamson M., Bauer B., Andreasson S.,

Query Match

Best Local Similarity

Matches 6; Conservative

54.4%; Score 37; DB 1; Length 2273;

Indels 2; Gaps 0;

1 CFAWQAM 8

779 CFAWQDM 786

RESULT 15

DYHB\_CHLRE

ID DYHB\_CHLRE STANDARD; PRT; 4568 AA.  
 AC Q39565;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DB Dynein beta chain, flagellar outer arm.  
 GN ODA4 OR ODA-4 OR SUP1.  
 OS Chlamydomonas reinhardtii.  
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
 OC Chlamydomonadales; Chlamydomonas.  
 OX NCBI\_TaxID=3055;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=21gr;  
 RX MEDLINE=94274778; PubMed=8006077;  
 RA Mitchell D.R., Brown K.S.,  
 RT "Sequence analysis of the Chlamydomonas alpha and beta dynein heavy  
 RT chain genes.",  
 RL J. Cell Sci. 107:635-644(1994).  
 CC -!- FUNCTION: FORCE GENERATING PROTEIN OF EUKARYOTIC CILIA AND  
 CC FLAGELLA. PRODUCES FORCE TOWARDS THE MINUS ENDS OF MICROTUBULES.  
 CC DYNEIN HAS ATPASE ACTIVITY.  
 CC -!- SUBUNIT: CONSISTS OF AT LEAST 3 HEAVY CHAINS (ALPHA, BETA AND  
 CC GAMMA), 2 INTERMEDIATE CHAINS AND 8 LIGHT CHAINS.  
 CC -!- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; U02963; AAA19956.1;  
 DR InterPro; IPR004273; Dynein heavy.  
 DR Pfam; PF03028; Dynein heavy; 1.  
 KW Motor protein; Microtubules; Dynein; ATP-binding; Flagella;  
 KM Coiled coil.  
 FT DOMAIN 277 293  
 FT COILED COIL (POTENTIAL).  
 FT DOMAIN 1158 1175  
 FT COILED COIL (POTENTIAL).  
 FT DOMAIN 1372 1400  
 FT COILED COIL (POTENTIAL).  
 FT DOMAIN 1614 1650  
 FT COILED COIL (POTENTIAL).  
 FT DOMAIN 1778 1825  
 FT COILED COIL (POTENTIAL).  
 FT DOMAIN 2017 2045  
 FT MICROTUBULE-BINDING (POTENTIAL).  
 FT DOMAIN 2831 2848  
 FT COILED COIL (POTENTIAL).  
 FT DOMAIN 3106 3162  
 FT COILED COIL (POTENTIAL).  
 FT DOMAIN 3648 3728  
 FT COILED COIL (POTENTIAL).  
 FT NP\_BIND 1919 1926  
 FT ATP (POTENTIAL).  
 FT NP\_BIND 2202 2209  
 FT ATP (POTENTIAL).  
 FT NP\_BIND 2530 2537  
 FT ATP (POTENTIAL).  
 FT NP\_BIND 2879 2886  
 FT ATP (POTENTIAL).  
 SQ SEQUENCE 4568 AA; 519961 MW; 9A9A5393C7C36AE7 CRC64;

Query Match

Best Local Similarity

Matches 5; Conservative

54.4%; Score 37; DB 1; Length 4568;

Indels 4; Gaps 0;

1 CFAWQAMKVR 12

1852 CFQWQSLRYIQ 1863

Search completed: February 21, 2003, 07:51:42

Job time : 6.2093 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:44:43 ; Search time 21.6744 Seconds  
(without alignments)  
114.078 Million cell updates/sec

Title: US-09-743-107b-95  
Perfect score: 68  
Sequence: 1 CFAWQRMKRV 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriaph:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	76.5	711	4 Q8TCD2	Q8tcd2 homo sapien
2	48	70.6	38	4 Q9UCY5	Q9ucy5 homo sapien
3	44	64.7	33	6 Q9TR80	Q9tr80 ovis aries
4	42	61.8	121	16 Q8U6K3	Q8u6k3 agrobacteri
5	40	58.8	114	11 Q83104	Q83104 rattus norv
6	40	58.8	234	11 Q83112	Q83112 rattus norv
7	39	57.4	511	16 Q8Z462	Q8z462 salmonella
8	39	57.4	1121	10 Q9SWS6	Q9sws6 lycopersico
9	38	55.9	134	2 Q87054	Q87054 vibrio chol
10	38	55.9	289	10 Q9C6N2	Q9c6n2 arabidopsis
11	38	55.9	329	16 Q98GW4	Q98gw4 rhizobium l
12	38	55.9	561	16 Q9KXK6	Q9kxk6 vibrio chol
13	38	55.9	759	4 Q9BVH6	Q9bvh6 homo sapien
14	38	55.9	1265	4 Q9P2G7	Q9p2g7 homo sapien
15	38	55.9	1417	4 Q9NYF4	Q9nyf4 homo sapien
16	38	55.9	1561	4 Q9BW93	Q9bw93 homo sapien

17	38	55.9	1787	4 Q9UPS0	Q9ups0 homo sapien
18	37	54.4	161	16 Q8U6L8	Q8u6l8 agrobacteri
19	37	54.4	202	16 Q83603	Q83603 treponema p
20	37	54.4	205	16 Q986A0	Q986a0 rhizobium l
21	37	54.4	206	16 Q8UHC2	Q8uhc2 agrobacteri
22	37	54.4	206	16 Q92RHS	Q92rhs rhizobium m
23	37	54.4	208	16 Q9YFK3	Q9yfk3 brucella me
24	37	54.4	233	2 Q9AJ90	Q9aj90 actinobacil
25	37	54.4	292	2 Q9JRR3	Q9jrr3 actinobacil
26	37	54.4	292	2 Q9AQB8	Q9aqb8 actinobacil
27	37	54.4	294	2 Q6G251	Q6g251 actinobacil
28	37	54.4	303	12 Q40640	Q40640 saimirine
29	37	54.4	303	12 Q92585	Q92585 saimirine
30	37	54.4	337	16 Q8UJL5	Q8ujl5 agrobacteri
31	37	54.4	408	17 Q8ZVY6	Q8zvy6 pyrobaculum
32	37	54.4	488	13 Q9DGD7	Q9dgd7 oryzias lat
33	37	54.4	494	16 Q9HY41	Q9hy41 pseudomonas
34	37	54.4	499	16 Q92GI5	Q92gi5 rickettsia
35	37	54.4	502	10 Q65237	Q65237 arabidopsis
36	37	54.4	584	17 Q9HPA3	Q9hpa3 halobacteri
37	37	54.4	676	5 Q9N4T5	Q9n4t5 caenorhabdi
38	37	54.4	676	10 Q9FG26	Q9fg26 arabidopsis
39	37	54.4	737	5 Q95QJ8	Q95qj8 caenorhabdi
40	37	54.4	830	10 Q945T7	Q945t7 hordeum vul
41	37	54.4	1137	10 Q9M7A9	Q9m7a9 oryza sativ
42	37	54.4	1139	10 Q8VWNL	Q8vwnl triticum ae
43	37	54.4	1836	16 Q9PEL7	Q9pel7 xyella fas
44	37	54.4	2310	11 Q35600	Q35600 mus musculu
45	37	54.4	2858	5 Q9GYP6	Q9gyp6 caenorhabdi

## ALIGNMENTS

## RESULT 1

Q8TCD2 ID Q8TCD2 PRELIMINARY; PRT; 711 AA.  
AC Q8TCD2;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Lactotransferrin.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PROSTATE;  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC022347; AAH22347.1; -  
SQ SEQUENCE 711 AA; 78327 MW; 1B9C7EE097C45FAF CRC64;

Query Match 76.5%; Score 52; DB 4; Length 711;  
Best Local Similarity 81.8%; Pred. NO. 0.49;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CFAWQRMKRV 11  
Db 39 CFWQRMKRV 49

## RESULT 2

Q9UCY5 ID Q9UCY5 PRELIMINARY; PRT; 38 AA.  
AC Q9UCY5;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE Lactoferrin homolog (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96081613; PubMed=8551695;  
 RA Sato I.;  
 RT "Characterization of the 84-kDa protein with ABH activity in human seminal plasma";  
 RL Jpn. J. Legal Med. 49:281-293(1995).  
 DR HSP; P02788; IBA.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 1.  
 SQ SEQUENCE 38 AA; 4459 MW; 0402F490B5E3DDEB CRC64;  
 Query Match 70.6%; Score 48; DB 4; Length 38;  
 Best Local Similarity 81.8%; Pred. No. 0.13;  
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 FAWQAMRKVR 12  
 | | | | |  
 Db 21 FQWQNNKVR 31  
 | | | | |  
 RESULT 3  
 Q9TR80 PRELIMINARY; PRT; 33 AA.  
 AC Q9TR80;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Lactoferrin (Fragment).  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=95127729; PubMed=7827104;  
 RA Qian Z.Y., Jolles P., Migliore-Samour D., Fiat A.M.;  
 RL Biochim. Biophys. Acta 1243:25-32(1995).  
 DR HSP; O77698; ICE2.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 1.  
 SQ SEQUENCE 33 AA; 3914 MW; D1904CAF15A73961 CRC64;  
 Query Match 64.7%; Score 44; DB 6; Length 33;  
 Best Local Similarity 54.5%; Pred. No. 0.56;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CFAWQAMRKV 11  
 | | | | |  
 Db 19 CYQWQKMKRL 29  
 | | | | |  
 RESULT 4  
 Q8U6K3 PRELIMINARY; PRT; 121 AA.  
 AC Q8U6K3;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Hypothetical protein At4g804.  
 GN ATU4804.  
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.  
 OX NCBI\_TaxID=176299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21608550; PubMed=11743193;  
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,  
 Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,

RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,  
 Chapman P., Ciendinning J., Deatherage G., Gillet W., Grant C.,  
 Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,  
 Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,  
 Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
 Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
 Chumley F., Tingey S.V., Tomb J.F., Gordon M.P., Olson M.V.,  
 Nester E.W.;  
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens C58";  
 RL Science 294:2317-2323(2001).  
 DR EMBL; A2009409; AAL45598.1; --  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 121 AA; 14085 MW; 5EAF1617A3CEA53 CRC64;

Query Match 61.8%; Score 42; DB 16; Length 121;  
 Best Local Similarity 63.6%; Pred. No. 4.8;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFAWQAMRKV 11  
 | | | | |  
 Db 14 CLAWQRNRV 24  
 | | | | |

RESULT 5  
 Q63104 PRELIMINARY; PRT; 114 AA.  
 AC Q63104;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Carcinoembryonic antigen (Fragment).  
 GN CEA4.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=BD II; TISSUE=LIVER;  
 RX MEDLINE=89214106; PubMed=2708349;  
 RA Kodellja V., Lucas K., Barnert S., von Kleist S., Thompson J.A.,  
 Zimmermann W.A.;  
 RT "Identification of a carcinoembryonic antigen gene family in the rat: Analysis of the N-terminal domains reveals immunoglobulin-like, hypervariable regions";  
 RT J. Biol. Chem. 264:6906-6912(1989).  
 RL EMBL; M60026; AAA40911.1; --  
 PT NON\_TER 1  
 SQ SEQUENCE 114 AA; 12832 MW; 3AE108689B061696 CRC64;

Query Match 58.8%; Score 40; DB 11; Length 114;  
 Best Local Similarity 60.0%; Pred. No. 10;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 FAWQAMRKV 11  
 | | | | |  
 Db 42 FAWYRGLRRI 51  
 | | | | |

RESULT 6  
 Q63112 PRELIMINARY; PRT; 234 AA.  
 AC Q63112;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Carcinoembryonic antigen-related protein (Fragment).  
 GN CGM4.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;

RN SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;  
 RA Rebeck S., Lucas K., Thompson J.A., Zimmermann W.A.;  
 RL Submitted (AUG-1990) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;  
 RX MEDLINE=90243655; PubMed=2335509;  
 RA Rebeck S., Lucas K., Thompson J.A., Zimmermann W.;  
 RT "cDNA and gene analyses imply a novel structure for a rat  
 carcinoembryonic antigen-related protein.";  
 RL J. Biol. Chem. 265:7872-7879(1990).  
 DR EMBL: M32475; AAA66038.1; --  
 DR InterPro; IPR003599; IG\_1  
 DR InterPro; IPR003599; IG\_2  
 DR Pfam; PF00047; IG\_1  
 DR Pfam; PF00047; IG\_2  
 DR SMART; SM00409; IG\_1  
 DR SMART; SM00409; IG\_2  
 FT NON TER 1  
 SQ SEQUENCE 234 AA; 26171 MW; 69B9C9BE0C773F2A CRC64;  
 Query Match 58.8%; Score 40; DB 11; Length 234;  
 Best Local Similarity 60.0%; Pred. No. 22;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 FAWQRMKRV 11  
 DB 42 FAWYRGRLKI 51  
 RESULT 7  
 ID Q82462 PRELIMINARY; PRT; 511 AA.  
 AC Q82462;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical protein STY3070.  
 GN STY3070.  
 OS Salmonella typhi.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Salmonella.  
 OX NCBI\_TaxID=601;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CT18;  
 RX MEDLINE=21534947; PubMed=11677608;  
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,  
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,  
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,  
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,  
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Garra P., Parry C.,  
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,  
 RA Whitehead S., Barrall B.G.;  
 RT "Complete genome sequence of a multiple drug resistant Salmonella  
 enterica serovar Typhi CT18.";  
 RL Nature 413:848-852(2001).  
 DR EMBL; AL627276; CAD06049.1; --  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 511 AA; 58136 MW; E2DD124E10D178B CRC64;  
 Query Match 57.4%; Score 39; DB 16; Length 511;  
 Best Local Similarity 58.3%; Pred. No. 72;  
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 CFAMQRMKRV 12  
 DB 350 CFAMQRMKRV 361  
 RESULT 8  
 Q9SWS6

ID Q9SWS6 PRELIMINARY; PRT; 1121 AA.  
 AC Q9SWS6;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Phytochrome B2.  
 GN PHYB2.  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; easterids I; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99413290; PubMed=10485280;  
 RA Kerckhoffs L.H., Kelmenson P.M., Schreuder M.E., Kendrick C.I.,  
 RA Kendrick R.E., Hanhart C.J., Koornneef M., Pratt L.H.,  
 RA Cordonnier-Pratt M.M.;  
 RT "Characterization of the gene encoding the apoprotein of phytochrome  
 B2 in tomato, and identification of molecular lesions in two mutant  
 alleles.";  
 RL Mol. Gen. Genet. 261:901-907(1999).  
 DR EMBL; AF122901; AAD50631.1; --  
 DR InterPro; IPR002198; ADH short.  
 DR InterPro; IPR003594; ATPBind\_ATPase.  
 DR InterPro; IPR003018; GAF.  
 DR InterPro; IPR003661; His\_kinA.  
 DR InterPro; IPR004359; HIS\_KIN\_sig.  
 DR InterPro; IPR001610; PAC.  
 DR InterPro; IPR000014; PAS domain.  
 DR InterPro; IPR001294; Phytochrome.  
 DR Pfam; PF01590; GAF; 1.  
 DR Pfam; PF02518; HATPase\_c; 1.  
 DR Pfam; PF00989; PAS; 2.  
 DR Pfam; PF00360; phytochrome; 1.  
 DR Pfam; PF00512; signal; 1.  
 DR PRINTS; PR01033; PHYTOCHROME.  
 DR SMART; SM00065; GAF; 1.  
 DR SMART; SM00387; HATPase\_c; 1.  
 DR SMART; SM00386; HSKA; 1.  
 DR SMART; SM00086; PAC; 1.  
 DR SMART; SM00091; PAS; 2.  
 DR TIGRFAMs; TIGR00229; sensory box; 2.  
 DR PROSITE; PS00061; ADH SHORT; UNKNOWN\_1.  
 DR PROSITE; PS00245; PHYTOCHROME\_1; 1.  
 DR PROSITE; PS00446; PHYTOCHROME\_2; 1.  
 SQ SEQUENCE 1121 AA; 125308 MW; ED9EDA704BB37F27 CRC64;  
 Query Match 57.4%; Score 39; DB 10; Length 1121;  
 Best Local Similarity 54.5%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 CFAMQRMKRV 11  
 DB 777 CFEMTAMEKL 787  
 RESULT 9  
 ID Q87054 PRELIMINARY; PRT; 134 AA.  
 AC Q87054;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE Z50F protein.  
 GN Z50F.  
 OS Vibrio cholerae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=666;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=Z17561;  
 RA Fallarino A.;

```

RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ231110; CAA13152.1; -- B340DBC6554BFAB CRC64;
SQ SEQUENCE 134 AA; 15496 MW; 1; Mismatches 3; Indels 0; Gaps 0;

Query Match 55.9%; Score 38; DB 2; Length 134;
Best Local Similarity 63.6%; Pred. No. 28;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 FAWQAMRKVR 12
DB 7 FAWRALRLKR 17

RESULT 10
Q9C6N2 PRELIMINARY; PRT; 289 AA.
AC Q9C6N2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Hypothetical 33.3 kDa protein.
GN F28L5.13.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Krenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzali A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana."
RL Nature 408:816-820(2000).
DR EMBL: AC079280; AAG50577.1; --
KW Hypothetical protein.
SQ SEQUENCE 289 AA; 33338 MW; 753AA27BED0F840C CRC64;

Query Match 55.9%; Score 38; DB 10; Length 289;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CFANQAMRKVR 12
DB 8 CFTWBYARVR 19

RESULT 11
Q98GW4 PRELIMINARY; PRT; 329 AA.
AC Q98GW4;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Membrane lipoprotein.
GN MLR3148.
OS Rhizobium loti (Mesorhizobium loti).
```

```

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsumoto A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL: AP003001; BA850102.1; --
DR InterPro; IPR003760; Bmp.
DR Pfam; PF02608; Bmp; 1.
KW Lipoprotein; Complete proteome.
SQ SEQUENCE 329 AA; 34754 MW; 78FC399867F6F3D2 CRC64;

Query Match 55.9%; Score 38; DB 16; Length 329;
Best Local Similarity 60.0%; Pred. No. 69;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 FAWQAMRKVR 11
DB 89 FAWEDALKKV 98

RESULT 12
Q9KKX6 PRELIMINARY; PRT; 561 AA.
AC Q9KKX6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Methyl-accepting chemotaxis protein.
GN VCA0974.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR Ni6961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
DR EMBL: AE004423; AAF96870.1; --
DR HSP; P02942; 1QU7.
DR TIGR; VCA0974; --
DR InterPro; IPR004089; Chmtaxis_trans.
DR InterPro; IPR003660; HAMP.
DR InterPro; IPR004090; Me_chemotaxis.
DR InterPro; IPR000727; T_SNARE.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF00015; MCPsignal; 1.
DR PRINTS; PR00260; CHEMTRNSDUCR.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00283; MA; 1.
KW Complete proteome.
SQ SEQUENCE 561 AA; 61113 MW; DDA599779A000D64 CRC64;

Query Match 55.9%; Score 38; DB 16; Length 561;
Best Local Similarity 63.6%; Pred. No. 1.2e+02;
```



Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 FAWORAMKVR 12  
 |||:|:|  
 Db 7 FAWMALRLKR 17

RESULT 13

Q9BVH6 PRELIMINARY; PRT; 759 AA.  
 ID Q9BVH6  
 AC Q9BVH6  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Hypothetical 84.7 kDa protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=EYE;  
 RA Strausberg R.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DDJ databases.  
 DR EMBL; BC001202; AAH01202.1; -.  
 DR InterPro; IPR003347; TF\_JmJC.  
 DR Pfam; PF02373; jmjC; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 759 AA; 84712 MW; 5A13F19C0EE952EF CRC64;

Query Match 55.9%; Score 38; DB 4; Length 759;  
 Best Local Similarity 70.0%; Pred. No. 1.6e+02;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 AWORAMKVR 12  
 |||:|:|  
 Db 17 AWKRAVRGR 26

RESULT 14

Q9P2G7 PRELIMINARY; PRT; 1265 AA.  
 ID Q9P2G7  
 AC Q9P2G7  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE KIAA1380 protein (Fragment).  
 GN KIAA1380.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRAIN;  
 RX MEDLINE=20181126; PubMed=10718198;  
 RA Nagase T., Kikuno R., Ishikawa K., Hirosewa M., Ohara O.;  
 RT Prediction of the coding sequences of unidentified human genes.XVI.  
 RT The complete sequences of 150 new cDNA clones from brain which code  
 RT for large proteins in vitro.;  
 RL DNA Res. 7:65-73(2000).  
 DR EMBL; AB037801; BAA92618.1; -.  
 DR InterPro; IPR003347; TF\_JmJC.  
 DR Pfam; PF02373; jmjC; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 1265 AA; 141291 MW; 33DB5BE53A7D9EBE CRC64;

Query Match 55.9%; Score 38; DB 4; Length 1265;  
 Best Local Similarity 70.0%; Pred. No. 2.8e+02;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 AWORAMKVR 12  
 |||:|:|

Db 560 AWKRAVRGR 569

RESULT 15

Q9NYF4 PRELIMINARY; PRT; 1417 AA.  
 ID Q9NYF4  
 AC Q9NYF4  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Putative zinc finger protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BONE MARROW;  
 RX MEDLINE=20541717; PubMed=11087669;  
 RA Lai F., Godley L.A., Fernald A.A., Orelli B.J., Pamintuan L., Zhao N.,  
 RA Le Beau M.W.;  
 RT "cDNA Cloning and Genomic Structure of Three Genes Localized to Human  
 RT Chromosome Band 5q31 Encoding Potential Nuclear Proteins.";  
 RL Genomics 70:123-130(2000).  
 DR EMBL; AF251039; AAF63765.1; -.  
 DR InterPro; IPR003347; TF\_JmJC.  
 DR Pfam; PF02373; jmjC; 1.  
 SQ SEQUENCE 1417 AA; 154720 MW; E9CF503D8654FD2B CRC64;

Query Match 55.9%; Score 38; DB 4; Length 1417;  
 Best Local Similarity 70.0%; Pred. No. 3.1e+02;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 AWORAMKVR 12  
 |||:|:|  
 Db 675 AWKRAVRGR 684

Search completed: February 21, 2003, 08:00:48  
 Job time : 22.6744 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:37:21 ; Search time 28.093 Seconds  
(without alignments)  
56.918 Million cell updates/sec

Title: US-09-743-107b-96

Perfect score: 64

Sequence: 1 CFQLKKNWKKYR 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A\_Geneseq\_101002.\*

```

1: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-embl/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-embl/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-embl/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-embl/AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	93.8	12	21	AA178096 Human lactoferrin
2	54	84.4	12	21	AA178080 Human lactoferrin
3	52	81.2	12	21	AA178091 Human lactoferrin
4	52	81.2	12	21	AA178092 Human lactoferrin
5	51	79.7	12	21	AA178081 Human lactoferrin
6	51	79.7	12	21	AA178083 Human lactoferrin
7	50	78.1	12	21	AA178097 Human lactoferrin
8	49	76.6	12	21	AA178071 Human lactoferrin
9	48	75.0	12	21	AA178038 Human lactoferrin
10	48	75.0	12	21	AA178046 Human lactoferrin

11	48	75.0	12	21	AA178047 Human lactoferrin
12	48	75.0	12	21	AA178084 Human lactoferrin
13	48	75.0	13	21	AA178037 Human lactoferrin
14	48	75.0	13	21	AA178048 Human lactoferrin
15	48	75.0	13	21	AA178049 Human lactoferrin
16	48	75.0	14	21	AA178036 Human lactoferrin
17	48	75.0	14	21	AA178050 Human lactoferrin
18	48	75.0	14	21	AA178051 Human lactoferrin
19	48	75.0	15	17	AA198554 Peptide for anti-u
20	48	75.0	15	17	AA178035 Human lactoferrin
21	48	75.0	15	21	AA178062 Human lactoferrin
22	48	75.0	15	21	AA178063 Human lactoferrin
23	48	75.0	16	21	AA178031 Human lactoferrin
24	48	75.0	16	21	AA178064 Human lactoferrin
25	48	75.0	16	21	AA178065 Human lactoferrin
26	48	75.0	17	21	AA178034 Human lactoferrin
27	48	75.0	17	21	AA178066 Human lactoferrin
28	48	75.0	17	21	AA178067 Human lactoferrin
29	48	75.0	17	21	AA178067 Human lactoferrin
30	48	75.0	18	15	AA19352 Human lactoferrin
31	48	75.0	18	17	AA13397 Advanced glycosyla
32	48	75.0	18	21	AA178033 Human lactoferrin
33	48	75.0	19	21	AA16867 Amino acid sequenc
34	48	75.0	19	21	AA178032 Human lactoferrin
35	48	75.0	20	13	AA121810 Anti microbial pep
36	48	75.0	20	14	AA14841 Lactoferrin-relate
37	48	75.0	20	15	AA148530 Lactoferrin derive
38	48	75.0	20	15	AA157461 Lactoferrin derive
39	48	75.0	20	15	AA157462 Lactoferrin derive
40	48	75.0	20	16	AA184698 Bovine lactoferrin
41	48	75.0	20	16	AA184699 Bovine lactoferrin
42	48	75.0	20	16	AA180263 Anti-parasitic lac
43	48	75.0	20	16	AA180264 Anti-parasitic lac
44	48	75.0	20	17	AA198553 Peptide for anti-u
45	48	75.0	20	17	AA191852 Lactoferrin-derive

#### ALIGNMENTS

RESULT 1  
AA178096  
ID AA178096 standard; Peptide; 12 AA.  
XX AC  
AA178096;  
XX AC  
DT 25-APR-2000 (first entry)  
XX DE  
Human lactoferrin derived peptide SEQ ID NO:96.  
XX Human; lactoferrin; modification; infection; inflammation; tumour;  
KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
KW urinary tract infection; colitis; Candida infection; fungicidal;  
KW bactericidal; preservative.  
XX OS Homo sapiens.  
OS Synthetic.  
XX WO200001730-A1.  
XX PD 13-JAN-2000.  
XX PF 06-JUL-1999; 99WO-SE01230.  
XX PR 06-JUL-1998; 98SE-0002441.  
XX PR 17-JUL-1998; 98SE-0002562.  
XX PR 29-DEC-1998; 98SE-0004614.  
XX PA (ASCI-) A+ SCI INVEST AB.  
XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
PT inflammations and tumors and for use in infant formula food -  
XX Claim 22; Page 38; 102pp; English.  
XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides are also  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.  
XX SQ Sequence 12 AA;  
Query Match 93.8%; Score 60; DB 21; Length 12;  
Best Local Similarity 91.7%; Pred. No. 0.00059;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CFOLKKNKKVR 12  
| | | | | | | | | | | |  
DB 1 CFOLKKNKKVR 12  
RESULT 2  
AAY78080  
ID AAY78080 standard; Peptide; 12 AA.  
XX AC AAY78080;  
XX DT 25-APR-2000 (first entry)  
XX DE Human lactoferrin derived peptide SEQ ID NO:80.  
XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
XX KW bactericidal; preservative.  
XX OS Homo sapiens.  
XX OS Synthetic.  
XX PN W0200001730-A1.  
XX PD 13-JAN-2000.  
XX PF 06-JUL-1999; 99WO-SE01230.  
XX PR 06-JUL-1998; 98SE-0002441.  
XX PR 17-JUL-1998; 98SE-0002562.  
XX PR 29-DEC-1998; 98SE-0004614.  
XX PA (ASCII-) A+ SCI INVEST AB.  
XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX PI WPI; 2000-147388/13.  
XX DR New peptides used for treatment and prevention of infections,  
XX PT inflammations and tumors and for use in infant formula food -  
XX PS Claim 22; Page 36; 102pp; English.  
XX AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides are also  
CC in food stuffs such as infant formula food. The peptides are also  
CC fungicidal and bactericidal and may also be used as preservatives.  
CC Even though native human lactoferrin have been shown to have desired  
CC anti-inflammatory anti-infectious and anti-tumoural properties they  
CC cannot be used clinically on a broad basis because of high production  
CC costs. Therefore, provision of peptides based on lactoferrin would  
CC enable them to be used for the same purposes as lactoferrin at lower  
CC cost.  
XX SQ Sequence 12 AA;  
Query Match 84.4%; Score 54; DB 21; Length 12;  
Best Local Similarity 75.0%; Pred. No. 0.0066;  
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CFOLKKNKKVR 12  
| | | | | | | | | | | |  
DB 1 CFOLKKNKKVR 12  
RESULT 3  
AAY78091  
ID AAY78091 standard; Peptide; 12 AA.  
XX AC AAY78091;  
XX DT 25-APR-2000 (first entry)  
XX DE Human lactoferrin derived peptide SEQ ID NO:91.  
XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
XX KW bactericidal; preservative.  
XX OS Homo sapiens.  
XX OS Synthetic.  
XX PN W0200001730-A1.  
XX PD 13-JAN-2000.  
XX PF 06-JUL-1999; 99WO-SE01230.  
XX PR 06-JUL-1998; 98SE-0002441.  
XX PR 17-JUL-1998; 98SE-0002562.  
XX PR 29-DEC-1998; 98SE-0004614.  
XX PA (ASCII-) A+ SCI INVEST AB.  
XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
XX PI WPI; 2000-147388/13.  
XX DR New peptides used for treatment and prevention of infections,  
XX PT inflammations and tumors and for use in infant formula food -  
XX PS Claim 22; Page 38; 102pp; English.  
XX AAY78001 to AAY78100 represent peptides having sequences based on human  
CC lactoferrin. The peptides are taken up in the intestine through  
CC binding to specific lactoferrin receptors and are then transported  
CC through the circulation. A medicinal product of the peptide or fragment  
CC can be used for treating and/or prevention of infections (such as  
CC urinary tract infections, colitis, and Candida infection on a mucosal  
CC membrane), inflammations and/or tumours. The peptides are also  
CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 12 AA;

Query Match 81.2%; Score 52; DB 21; Length 12;  
 Best Local Similarity 75.0%; Pred. No. 0.015;  
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQLKKNKKVR 12  
 ||| |.:|:|  
 Db 1 CFQWKNNKKVR 12

RESULT 4  
 AAY78092  
 ID AAY78092 standard; Peptide; 12 AA.

XX AC AAY78092;  
 XX DT 25-APR-2000 (first entry)  
 XX DE Human lactoferrin derived peptide SEQ ID NO:92.  
 XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
 XX KW bactericidal; preservative.

XX OS Homo sapiens.  
 XX OS Synthetic.  
 XX PN WO200001730-A1.  
 XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.  
 XX PR 06-JUL-1998; 98SE-0002441.  
 XX PR 17-JUL-1998; 98SE-0002562.  
 XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX PS WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 XX PT inflammations and tumors and for use in infant formula food -

XX PS Claim 22; Page 38; 102pp; English.

XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 XX CC lactoferrin. The peptides are taken up in the intestine through  
 XX CC binding to specific lactoferrin receptors and are then transported  
 XX CC through the circulation. A medicinal product of the peptide or fragment  
 XX CC can be used for treating and/or prevention of infections (such as  
 XX CC urinary tract infections, colitis, and Candida infection on a mucosal  
 XX CC membrane), inflammations and/or tumors. The peptides can also be used  
 XX CC in food stuffs such as infant formula food. The peptides are also  
 XX CC fungicidal and bactericidal and may also be used as preservatives.  
 XX CC Even though native human lactoferrin have been shown to have desired  
 XX CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 XX CC cannot be used clinically on a broad basis because of high production  
 XX CC costs. Therefore, provision of peptides based on lactoferrin would  
 XX CC enable them to be used for the same purposes as lactoferrin at lower  
 XX CC cost.

XX SQ Sequence 12 AA;

Query Match 81.2%; Score 52; DB 21; Length 12;  
 Best Local Similarity 75.0%; Pred. No. 0.015;  
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFQLKKNKKVR 12  
 ||| |.:|:|  
 Db 1 CFQWKNNKKVR 12

RESULT 5  
 AAY78081  
 ID AAY78081 standard; Peptide; 12 AA.

XX AC AAY78081;  
 XX DT 25-APR-2000 (first entry)  
 XX DE Human lactoferrin derived peptide SEQ ID NO:81.  
 XX KW Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 XX KW urinary tract infection; colitis; Candida infection; fungicidal;  
 XX KW bactericidal; preservative.

XX OS Homo sapiens.  
 XX OS Synthetic.

XX PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.  
 XX PR 06-JUL-1998; 98SE-0002441.  
 XX PR 17-JUL-1998; 98SE-0002562.  
 XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX PS WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,  
 XX PT inflammations and tumors and for use in infant formula food -

XX PS Claim 22; Page 36; 102pp; English.

XX CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 XX CC lactoferrin. The peptides are taken up in the intestine through  
 XX CC binding to specific lactoferrin receptors and are then transported  
 XX CC through the circulation. A medicinal product of the peptide or fragment  
 XX CC can be used for treating and/or prevention of infections (such as  
 XX CC urinary tract infections, colitis, and Candida infection on a mucosal  
 XX CC membrane), inflammations and/or tumors. The peptides can also be used  
 XX CC in food stuffs such as infant formula food. The peptides are also  
 XX CC fungicidal and bactericidal and may also be used as preservatives.  
 XX CC Even though native human lactoferrin have been shown to have desired  
 XX CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 XX CC cannot be used clinically on a broad basis because of high production  
 XX CC costs. Therefore, provision of peptides based on lactoferrin would  
 XX CC enable them to be used for the same purposes as lactoferrin at lower  
 XX CC cost.

XX SQ Sequence 12 AA;

Query Match 79.7%; Score 51; DB 21; Length 12;  
 Best Local Similarity 75.0%; Pred. No. 0.022;  
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFOLKKNKKVR 12  
 DB 1 CFQWQKNKKVR 12

## RESULT 6

AA78083  
 ID AAY78083 standard; Peptide; 12 AA.

AC AAY78083;

DT 25-APR-2000 (first entry)

DE Human lactoferrin derived peptide SEQ ID NO:83.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.

OS Synthetic.

PN WO200001730-A1.

XX 13-JAN-2000.

PF 06-JUL-1999; 99WO-SE01230.

PR 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

DR WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

PS Claim 22; Page 36; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 12 AA;

Query Match 79.7%; Score 51; DB 21; Length 12;

Best Local Similarity 75.0%; Pred. No. 0.022; 1; Indels 0; Gaps 0;  
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFOLKKNKKVR 12

DB 1 CFQWQKNKKVR 12

## RESULT 7

AA78097

ID AAY78097 standard; Peptide; 12 AA.

XX AAY78097;

DT 25-APR-2000 (first entry)

DE Human lactoferrin derived peptide SEQ ID NO:97.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.

XX Homo sapiens.

OS Synthetic.

PN WO200001730-A1.

XX 13-JAN-2000.

PF 06-JUL-1999; 99WO-SE01230.

PR 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

DR WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -

PS Claim 19; Page 94; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX Sequence 12 AA;

Query Match 78.1%; Score 50; DB 21; Length 12;

Best Local Similarity 83.3%; Pred. No. 0.033; 2; Indels 0; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFOLKKNKKVR 12

DB 1 CFALKKKKKKVR 12

## RESULT 8

AA78071

ID AAY78071 standard; Peptide; 12 AA.

XX AAY78071;

DT 25-APR-2000 (first entry)

DE Human lactoferrin derived peptide SEQ ID NO:71.

XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 XX Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200001730-A1.  
 XX  
 XX 13-JAN-2000.  
 XX  
 XX 06-JUL-1999; 99WO-SE01230.  
 XX  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 XX (ASCI-) A+ SCI INVEST AB.  
 XX  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 XX  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 PT  
 PS Claim 22; Page 35; 102pp; English.  
 XX  
 CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 XX Sequence 12 AA;  
 SQ  
 Query Match 76.6%; Score 49; DB 21; Length 12;  
 Best Local Similarity 66.7%; Pred. No. 0.049;  
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQLKKNMKVVR 12  
 DB 1 CFQQAQRNMRKVR 12  
 RESULT 9  
 AAY78038  
 ID AAY78038 standard; Peptide; 12 AA.  
 XX  
 XX AAY78038;  
 AC  
 XX 25-APR-2000 (first entry)  
 DT  
 XX Human lactoferrin derived peptide SEQ ID NO:38.  
 DE  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 XX Homo sapiens.  
 OS

OS Synthetic.  
 XX  
 PN WO200001730-A1.  
 XX  
 PD 13-JAN-2000.  
 XX  
 XX 06-JUL-1999; 99WO-SE01230.  
 PF  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX  
 XX (ASCI-) A+ SCI INVEST AB.  
 PA  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 PI WPI; 2000-147388/13.  
 XX  
 XX New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 PT  
 PS Claim 12; Page 70; 102pp; English.  
 XX  
 CC AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumours. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX  
 XX Sequence 12 AA;  
 SQ  
 Query Match 75.0%; Score 48; DB 21; Length 12;  
 Best Local Similarity 66.7%; Pred. No. 0.073;  
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQLKKNMKVVR 12  
 DB 1 CFQQAQRNMRKVR 12  
 RESULT 10  
 AAY78046  
 ID AAY78046 standard; Peptide; 12 AA.  
 XX  
 XX AAY78046;  
 AC  
 XX 25-APR-2000 (first entry)  
 DT  
 XX Human lactoferrin derived peptide SEQ ID NO:46.  
 DE  
 XX Human; lactoferrin; modification; infection; inflammation; tumour;  
 KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX  
 XX Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200001730-A1.  
 XX  
 XX 13-JAN-2000.  
 PD  
 XX 06-JUL-1999; 99WO-SE01230.  
 PF

XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 DR New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 15; Page 35; 102pp; English.  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX SQ Sequence 12 AA;  
 Query Match 75.0%; Score 48; DB 21; Length 12;  
 Best Local Similarity 66.7%; Pred. No. 0.073; 1; Indels 0; Gaps 0;  
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQLKKNMKVVR 12  
 DB ||| :||:|  
 1 CFQWQNNMKVVR 12  
 RESULT 11  
 AAY78047  
 ID AAY78047 standard; Peptide; 12 AA.  
 AC AAY78047;  
 XX 25-APR-2000 (first entry)  
 DT Human lactoferrin derived peptide SEQ ID NO:47.  
 DE Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200001730-A1.  
 PN 13-JAN-2000.  
 XX 06-JUL-1998; 98WO-SE01230.  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 DR New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX

PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 DR New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX Claim 18; Page 73; 102pp; English.  
 XX AAY78001 to AAY78100 represent peptides having sequences based on human  
 CC lactoferrin. The peptides are taken up in the intestine through  
 CC binding to specific lactoferrin receptors and are then transported  
 CC through the circulation. A medicinal product of the peptide or fragment  
 CC can be used for treating and/or prevention of infections (such as  
 CC urinary tract infections, colitis, and Candida infection on a mucosal  
 CC membrane), inflammations and/or tumors. The peptides can also be used  
 CC in food stuffs such as infant formula food. The peptides are also  
 CC fungicidal and bactericidal and may also be used as preservatives.  
 CC Even though native human lactoferrin have been shown to have desired  
 CC anti-inflammatory anti-infectious and anti-tumoural properties they  
 CC cannot be used clinically on a broad basis because of high production  
 CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.  
 XX SQ Sequence 12 AA;  
 Query Match 75.0%; Score 48; DB 21; Length 12;  
 Best Local Similarity 66.7%; Pred. No. 0.073; 1; Indels 0; Gaps 0;  
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 CFQLKKNMKVVR 12  
 DB ||| :||:|  
 1 CFQWQNNMKVVR 12  
 RESULT 12  
 AAY78084  
 ID AAY78084 standard; Peptide; 12 AA.  
 AC AAY78084;  
 XX 25-APR-2000 (first entry)  
 DT Human lactoferrin derived peptide SEQ ID NO:84.  
 DE Human; lactoferrin; modification; infection; inflammation; tumour;  
 XX food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;  
 KW urinary tract infection; colitis; Candida infection; fungicidal;  
 KW bactericidal; preservative.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO200001730-A1.  
 PN 13-JAN-2000.  
 XX 06-JUL-1998; 98WO-SE01230.  
 XX 06-JUL-1998; 98SE-0002441.  
 PR 17-JUL-1998; 98SE-0002562.  
 PR 29-DEC-1998; 98SE-0004614.  
 XX (ASCI-) A+ SCI INVEST AB.  
 XX Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;  
 XX WPI; 2000-147388/13.  
 DR New peptides used for treatment and prevention of infections,  
 PT inflammations and tumors and for use in infant formula food -  
 XX

PS Claim 22; Page 36; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through

CC binding to specific lactoferrin receptors and are then transported

CC through the circulation. A medicinal product of the peptide or fragment

CC can be used for treating and/or prevention of infections (such as

CC urinary tract infections, colitis, and Candida infection on a mucosal

CC membrane), inflammations and/or tumours. The peptides can also be used

CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.

CC Even though native human lactoferrin have been shown to have desired

CC anti-inflammatory anti-infectious and anti-tumoural properties they

CC cannot be used clinically on a broad basis because of high production

CC costs. Therefore, provision of peptides based on lactoferrin would

CC enable them to be used for the same purposes as lactoferrin at lower

CC cost.

XX Sequence 12 AA;

SQ Query Match 75.0%; Score 48; DB 21; Length 12;

Best Local Similarity 66.7%; Pred. No. 0.073;

Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQLKKNMKVVR 12

DB 1 CFQWRNMRKVR 12

RESULT 13

AAY78037

ID AAY78037 standard; Peptide; 13 AA.

XX AC AAY78037;

XX DT 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:37.

XX Human; lactoferrin; modification; infection; inflammation; tumour;

KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;

KW urinary tract infection; colitis; Candida infection; fungicidal;

KW bactericidal; preservative.

XX Homo sapiens.

OS Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Matteby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,

PT inflammations and tumors and for use in infant formula food -

PS Claim 12; Page 70; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through

CC binding to specific lactoferrin receptors and are then transported

CC through the circulation. A medicinal product of the peptide or fragment

CC can be used for treating and/or prevention of infections (such as

CC urinary tract infections, colitis, and Candida infection on a mucosal

CC membrane), inflammations and/or tumours. The peptides can also be used

CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.

CC Even though native human lactoferrin have been shown to have desired

CC anti-inflammatory anti-infectious and anti-tumoural properties they

CC cannot be used clinically on a broad basis because of high production

CC costs. Therefore, provision of peptides based on lactoferrin would

CC enable them to be used for the same purposes as lactoferrin at lower

CC cost.

XX Sequence 13 AA;

SQ Query Match 75.0%; Score 48; DB 21; Length 13;

Best Local Similarity 66.7%; Pred. No. 0.079;

Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQLKKNMKVVR 12

DB 2 CFQWRNMRKVR 13

RESULT 14

AAY78048

ID AAY78048 standard; Peptide; 13 AA.

XX AC AAY78048;

XX DT 25-APR-2000 (first entry)

XX Human lactoferrin derived peptide SEQ ID NO:48.

XX Human; lactoferrin; modification; infection; inflammation; tumour;

KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;

KW urinary tract infection; colitis; Candida infection; fungicidal;

KW bactericidal; preservative.

XX Homo sapiens.

OS Synthetic.

XX WO200001730-A1.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-SE01230.

XX 06-JUL-1998; 98SE-0002441.

PR 17-JUL-1998; 98SE-0002562.

PR 29-DEC-1998; 98SE-0004614.

XX (ASCI-) A+ SCI INVEST AB.

XX Hanson LA, Matteby-Baltzer I, Baltzer L, Dolphin GT;

XX WPI; 2000-147388/13.

XX New peptides used for treatment and prevention of infections,

PT inflammations and tumors and for use in infant formula food -

PS Claim 15; Page 74; 102pp; English.

XX AAY78001 to AAY78100 represent peptides having sequences based on human

CC lactoferrin. The peptides are taken up in the intestine through

CC binding to specific lactoferrin receptors and are then transported

CC through the circulation. A medicinal product of the peptide or fragment

CC can be used for treating and/or prevention of infections (such as

CC urinary tract infections, colitis, and Candida infection on a mucosal

CC membrane), inflammations and/or tumours. The peptides can also be used

CC in food stuffs such as infant formula food. The peptides are also

CC fungicidal and bactericidal and may also be used as preservatives.

CC Even though native human lactoferrin have been shown to have desired

CC anti-inflammatory anti-infectious and anti-tumoural properties they

CC cannot be used clinically on a broad basis because of high production



CC costs. Therefore, provision of peptides based on lactoferrin would  
 CC enable them to be used for the same purposes as lactoferrin at lower  
 CC cost.

XX SQ Sequence 13 AA;

Query Match 75.0%; Score 48; DB 21; Length 13;

Best Local Similarity 66.7%; Pred. No. 0.079;

Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFOLKKNKKVR 12

DB 2 CFQWRNKKVR 13

# RESULT 15

AAV78049

ID AAV78049 standard; Peptide; 13 AA.

XX AC AAV78049;

XX DT 25-APR-2000 (first entry)

XX DE Human lactoferrin derived peptide SEQ ID NO:49.

XX KW Human; lactoferrin; modification; infection; inflammation; tumour;

XX KW food; infant formula; anti-inflammatory; anti-microbial; anti-tumour;

XX KW urinary tract infection; colitis; Candida infection; fungicidal;

XX KW bactericidal; preservative.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200001730-A1.

XX PD 13-JAN-2000.

XX PF 06-JUL-1999; 99WO-SE01230.

XX PR 06-JUL-1998; 98SE-0002441.

XX PR 17-JUL-1998; 98SE-0002562.

XX PR 29-DEC-1998; 98SE-0004614.

XX PA (ASCI-) A+ SCI INVEST AB.

XX PI Hanson LA, Mattsby-Baltzer I, Baltzer L, Dolphin GT;

XX DR WPI; 2000-147388/13.

XX PT New peptides used for treatment and prevention of infections,

XX PT inflammations and tumors and for use in infant formula food -

XX PS Claim 18; Page 74; 102pp; English.

XX CC AAV78001 to AAV78100 represent peptides having sequences based on human

XX CC lactoferrin. The peptides are taken up in the intestine through

XX CC binding to specific lactoferrin receptors and are then transported

XX CC through the circulation. A medicinal product of the peptide or fragment

XX CC can be used for treating and/or prevention of infections (such as

XX CC urinary tract infections, colitis, and Candida infection on a mucosal

XX CC membrane), inflammations and/or tumours. The peptides can also be used

XX CC in food stuffs such as infant formula food. The peptides are also

XX CC fungicidal and bactericidal and may also be used as preservatives.

XX CC Even though native human lactoferrin have been shown to have desired

XX CC anti-inflammatory anti-infectious and anti-tumoural properties they

XX CC cannot be used clinically on a broad basis because of high production

XX CC costs. Therefore, provision of peptides based on lactoferrin would

XX CC enable them to be used for the same purposes as lactoferrin at lower

XX CC cost.

XX SQ Sequence 13 AA;

Query Match 75.0%; Score 48; DB 21; Length 13;

Best Local Similarity 66.7%; Pred. No. 0.079;  
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFOLKKNKKVR 12

DB 2 CFQWRNKKVR 13

Search completed: February 21, 2003, 07:56:45

Job time : 28.093 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 07:50:40 , Search time 8.93023 Seconds  
(without alignments)  
39.537 Million cell updates/sec

Title: US-09-743-107B-96  
Perfect score: 64  
Sequence: 1 CFOLKKNMKVR 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/aaa/5A\_COMB.psp:\*  
2: /cgn2\_6/prodata/1/aaa/5B\_COMB.psp:\*  
3: /cgn2\_6/prodata/1/aaa/6A\_COMB.psp:\*  
4: /cgn2\_6/prodata/1/aaa/6B\_COMB.psp:\*  
5: /cgn2\_6/prodata/1/aaa/PCUS\_COMB.psp:\*  
6: /cgn2\_6/prodata/1/aaa/backfiles1.psp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	75.0	18	1	US-08-204-487-3
2	48	75.0	18	2	US-08-485-948-8
3	48	75.0	18	2	US-08-628-380-8
4	48	75.0	18	2	US-08-475-055-8
5	48	75.0	20	1	US-07-753-161A-3
6	48	75.0	20	1	US-07-891-174-3
7	48	75.0	20	1	US-08-204-487-1
8	48	75.0	20	1	US-08-256-771-24
9	48	75.0	20	1	US-08-256-771-25
10	48	75.0	20	1	US-08-381-984-24
11	48	75.0	20	1	US-08-381-984-25
12	48	75.0	22	4	US-09-508-734-4
13	48	75.0	24	4	US-09-508-734-6
14	48	75.0	25	1	US-07-755-161A-10
15	48	75.0	25	1	US-07-891-174-10
16	48	75.0	25	1	US-08-204-487-7
17	48	75.0	29	4	US-09-508-734-8
18	48	75.0	36	1	US-07-753-161A-8
19	48	75.0	36	1	US-07-891-174-8
20	48	75.0	36	1	US-08-256-771-30
21	48	75.0	36	1	US-08-381-984-29
22	48	75.0	47	2	US-08-464-182A-6
23	48	75.0	47	2	US-08-406-271-6
24	48	75.0	50	2	US-08-693-274A-7
25	48	75.0	52	4	US-09-017-043A-3
26	48	75.0	53	2	US-08-464-182A-5
27	48	75.0	53	2	US-08-406-271-5

28 48 75.0 54 2 US-08-464-182A-2 Appli  
29 48 75.0 54 2 US-08-406-271-2 Appli  
30 48 75.0 694 3 US-08-724-586-2 Appli  
31 48 75.0 694 4 US-09-421-632-2 Appli  
32 48 75.0 694 4 US-09-932-190-2 Appli  
33 48 75.0 705 2 US-08-655-640-2 Appli  
34 48 75.0 708 2 US-08-655-640-4 Appli  
35 48 75.0 711 1 US-08-154-019-4 Appli  
36 48 75.0 711 1 US-08-461-333-4 Appli  
37 48 75.0 711 3 US-08-464-167-4 Appli  
38 48 75.0 711 3 US-09-158-313-4 Appli  
39 48 75.0 711 4 US-08-475-798-4 Appli  
40 45 70.3 711 1 US-08-145-681-2 Appli  
41 45 70.3 711 1 US-08-250-308-2 Appli  
42 45 70.3 711 1 US-08-453-703-2 Appli  
43 45 70.3 711 3 US-08-456-106-2 Appli  
44 45 70.3 711 3 US-08-456-108-2 Appli  
45 45 70.3 711 4 US-09-265-577-2 Appli

## ALIGNMENTS

RESULT 1  
US-08-204-487-3  
; Sequence 3, Application US/08204487  
; Patent No. 5585425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAOKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGEKI  
; APPLICANT: DOSAKO, SHUN'ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESS: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FCN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..18  
; OTHER INFORMATION: /note= "HUMAN LACTOFERRIN PEPTIDE"

OTHER INFORMATION: (20-37)"  
US-08-204-487-3

Query Match 75.0%; Score 48; DB 1; Length 18;  
Best Local Similarity 66.7%; Pred. No. 0.056;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFOLKKNMKVR 12  
||| :||:|  
Db 1 CFQWRNMRKVR 12

## RESULT 2

US-08-485-948-8  
; Sequence 8, Application US/08485948  
; Patent No. 585882  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,948  
; FILING DATE:  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/488,217  
; FILING DATE: JUNE 7, 1995  
; APPLICATION NUMBER: 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LF-Cl, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal

Query Match 75.0%; Score 48; DB 2; Length 18;  
Best Local Similarity 66.7%; Pred. No. 0.056;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFOLKKNMKVR 12  
||| :||:|  
Db 1 CFQWRNMRKVR 12

## RESULT 3

US-08-628-380-8  
; Sequence 8, Application US/08628380  
; Patent No. 5891341  
; GENERAL INFORMATION:  
; APPLICANT: LI, YONG MING  
; APPLICANT: VLASSARA, HELEN  
; APPLICANT: CERAMI, ANTHONY  
; TITLE OF INVENTION: AGE-MEDIATED INHIBITION OF ANTIBACTERIAL PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/628,380  
; FILING DATE: April 4, 1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/418,642  
; FILING DATE: APRIL 7, 1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 947-1-008 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; DESCRIPTION: LF-Cl, 8-25  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: internal

Query Match 75.0%; Score 48; DB 2; Length 18;  
Best Local Similarity 66.7%; Pred. No. 0.056;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFOLKKNMKVR 12  
||| :||:|  
Db 1 CFQWRNMRKVR 12

## RESULT 4

US-08-475-055-8  
; Sequence 8, Application US/08475055  
; Patent No. 5962245  
; GENERAL INFORMATION:  
; APPLICANT: YONG MING LI  
; APPLICANT: HELEN VLASSARA  
; APPLICANT: ANTHONY CERAMI  
; TITLE OF INVENTION: AGENTS FOR BINDING TO ADVANCED GLYCOSYLATION  
; TITLE OF INVENTION: ENDPRODUCTS, AND METHODS OF THEIR USE  
; NUMBER OF SEQUENCES: 9

```

;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,055
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/485,948
; FILING DATE:
; APPLICATION NUMBER: 08/488,217
; FILING DATE: JUNE 7, 1995
; APPLICATION NUMBER: 08/418,642
; FILING DATE: APRIL 7, 1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 947-1-008A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
;
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; DESCRIPTION: LF-CI, 8-25
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
;
; US-08-475-055-8
;
; Query Match 75.0%; Score 48; DB 2; Length 18;
; Best Local Similarity 66.7%; Pred No. 0.056;
; Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
;
; QY 1 CFQLKKNMKVR 12
; |||::|||
; Db 1 CFQQRNMRKVR 12
;
; RESULT 5
; US-07-755-161A-3
; Sequence 3, Application US/07755161A
; Patent No. 5304633
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptides and an
; TITLE OF INVENTION: Antimicrobial Agent
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb
; COMPUTER: IBM Compatible
;
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: DisplayWrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/755.161A
; FILING DATE: 19910905
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX: 202-371-8856
; TELEX:
;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 2
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 19"
; FEATURE:
; NAME/KEY: modified site
; LOCATION: 19
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "thiol group of
; OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with
; OTHER INFORMATION: thiol group of Cys residue at location 2"
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-07-755-161A-3
;
; Query Match 75.0%; Score 48; DB 1; Length 20;
```

Best Local Similarity 65.7%; Pred. No. 0.062; Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFOLKKMKVKR 12  
DB 2 CFQWRNRRKVR 13

RESULT 6  
US-07-891-174-3  
; Sequence 3, Application US/07891174  
; Patent No. 5317084  
; GENERAL INFORMATION:  
; APPLICANT: Mamoru TOMITA et al.  
; TITLE OF INVENTION: Antimicrobial peptides and an  
; TITLE OF INVENTION: Antimicrobial Agent  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Displaywrite  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/891,174  
; FILING DATE: 29-MAY-1992  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/755,161  
; FILING DATE: 05-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELEPHONE: 202-371-8850  
; TELEFAX: 202-371-8856  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; HYPOTHEICAL:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; ORGANISM:  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:  
; CELL LINE:  
; ORGANELLE:  
; IMMEDIATE SOURCE:  
; LIBRARY:  
; CLONE:  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT:  
; MAP POSITION:  
; UNITS:  
; FEATURE:  
; NAME/KEY: modified site

LOCATION: 2  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 2 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 19"  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 19  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 19 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 2"  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-891-174-3

Query Match 75.0%; Score 48; DB 1; Length 20;  
Best Local Similarity 66.7%; Pred. No. 0.062;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFOLKKMKVKR 12  
DB 2 CFQWRNRRKVR 13

RESULT 7  
US-08-204-487-1  
; Sequence 1, Application US/08204487  
; Patent No. 5565425  
; GENERAL INFORMATION:  
; APPLICANT: YAMAMOTO, NAOKI  
; APPLICANT: NAKASHIMA, HIDEKI  
; APPLICANT: MOSUCHI, WATARU  
; APPLICANT: TANAKA, SHIGERAKI  
; APPLICANT: DOSAKO, SHUN'ICHI  
; APPLICANT: KAWASAKI, YOSHIHIRO  
; APPLICANT: UCHIDA, TOSHIAKI  
; TITLE OF INVENTION: VIRAL INFECTION AND PROLIFERATION  
; TITLE OF INVENTION: INHIBITORS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &  
; ADDRESSEE: THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/204,487  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, PAULA A.  
; REGISTRATION NUMBER: 32,503  
; REFERENCE/DOCKET NUMBER: FUN-019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000

```
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..20
OTHER INFORMATION: /note= "ANTIBACTERIAL PEPTIDE
DERIVED FROM HUMAN LACTOFERRIN"
US-08-204-487-1
Query Match 75.0%; Score 48; DB 1; Length 20;
Best Local Similarity 66.7%; Pred. No. 0.062;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 CFQLKKNMKVR 12
DB 2 CFQWRNMRKVR 13
RESULT 8
US-08-256-771-24
; Sequence 24, Application US/08256771
; Patent No. 5656591
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING
; TITLE OF INVENTION: PRODUCTS THEREWITH
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,771
; FILING DATE: July 22, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "Cys residues are linked by
; OTHER INFORMATION: disulfide bond"
```

```
US-08-256-771-24
Query Match 75.0%; Score 48; DB 1; Length 20;
Best Local Similarity 66.7%; Pred. No. 0.062;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 CFQLKKNMKVR 12
DB 2 CFQWRNMRKVR 13
RESULT 9
US-08-256-771-25
; Sequence 25, Application US/08256771
; Patent No. 5656591
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING
; TITLE OF INVENTION: PRODUCTS THEREWITH
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,771
; FILING DATE: July 22, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note= "Cys residues are protected to
; OTHER INFORMATION: prevent disulfide bond"
US-08-256-771-25
Query Match 75.0%; Score 48; DB 1; Length 20;
Best Local Similarity 66.7%; Pred. No. 0.062;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 CFQLKKNMKVR 12
DB 2 CFQWRNMRKVR 13
RESULT 10
US-08-381-984-24
```

Sequence 24, Application US/08381984  
Patent No. 5804555  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIOXIDANT  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/381,984  
FILING DATE: April 11, 1995  
CLASSIFICATION: 252  
PRIOR APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "cysteine residues at positions 2  
OTHER INFORMATION: and 19 are bonded by disulfide linkage"  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "the specified peptide as well as  
OTHER INFORMATION: peptides including the specified peptide as a fragment thereof  
US-08-381-984-24  
Query Match 75.0%; Score 48; DB 1; Length 20;  
Best Local Similarity 66.7%; Pred. No. 0.062;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CFQKKNMKV 12  
DB 2 CFQWQNMKVR 13  
RESULT 11  
US-08-381-984-25  
Sequence 25, Application US/08381984  
Patent No. 5804555  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: ANTIOXIDANT  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack

STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/381,984  
FILING DATE: April 11, 1995  
CLASSIFICATION: 252  
PRIOR APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "the specified peptide as well as  
OTHER INFORMATION: peptides including the specified peptide as a fragment thereof  
US-08-381-984-25  
Query Match 75.0%; Score 48; DB 1; Length 20;  
Best Local Similarity 66.7%; Pred. No. 0.062;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CFQKKNMKV 12  
DB 2 CFQWQNMKVR 13  
RESULT 12  
US-09-508-734-4  
Sequence 4, Application US/09508734  
Patent No. 6423509  
GENERAL INFORMATION:  
APPLICANT: Samyang Genex Corporation  
TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and  
FILE REFERENCE: PA/SYG/00139  
CURRENT APPLICATION NUMBER: US/09/508,734  
CURRENT FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: PCT/KR99/00373  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: KR1998-29351  
PRIOR FILING DATE: 1998-07-13  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Kopatentin 1.71  
SEQ ID NO 4

```
LENGTH: 22
TYPE: PRT
ORGANISM: Homo sapiens
IS-09-508-734-4
Query Match 75.0%; Score 48; DB 4; Length 22;
Best Local Similarity 66.7%; Pred. No. 0.068;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

y 1 CFOLKKNMKVYR 12
   ||| :|||
   ||| :|||
b 2 CFQWRNMRKVR 13

RESULT 13
IS-09-508-734-6
Sequence 6, Application US/09508734
Patent No. 6423509
GENERAL INFORMATION:
APPLICANT: Samyang Genex Corporation
TITLE OF INVENTION: Mass production method of lactoferrin polypeptide from yeast and
TITLE OF INVENTION: useful microorganism thereof
FILE REFERENCE: PA/SYG/00139
CURRENT APPLICATION NUMBER: US/09/508,734
CURRENT FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/KR99/00373
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: KR1998-29351
PRIOR FILING DATE: 1998-07-13
NUMBER OF SEQ ID NOS: 12
SOFTWARE: KopatentIn 1.71
SEQ ID NO 6
LENGTH: 24
TYPE: PRT
ORGANISM: Homo sapiens
IS-09-508-734-6
Query Match 75.0%; Score 48; DB 4; Length 24;
Best Local Similarity 66.7%; Pred. No. 0.074;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

y 1 CFOLKKNMKVYR 12
   ||| :|||
   ||| :|||
b 3 CFQWRNMRKVR 14

RESULT 14
IS-07-755-161A-10
Sequence 10, Application US/07755161A
Patent No. 5304633
GENERAL INFORMATION:
APPLICANT: Mamoru TOMITA et al.
TITLE OF INVENTION: Antimicrobial Peptides and an
TITLE OF INVENTION: Antimicrobial Agent
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500KB
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/755,161A
FILING DATE: 19910905
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:
NAME: Warren M. Cheek Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELEPHONE: 202-371-8850
TELEFAX: 202-371-8856
TELEX:
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY: modified site
LOCATION: 4
IDENTIFICATION METHOD:
OTHER INFORMATION: /note= "thiol group of
OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with
OTHER INFORMATION: thiol group of Cys residue at location 21"
FEATURE:
NAME/KEY: modified site
LOCATION: 21
IDENTIFICATION METHOD:
OTHER INFORMATION: /note= "thiol group of
OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with
OTHER INFORMATION: thiol group of Cys residue at location 4"
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-07-755-161A-10
Query Match 75.0%; Score 48; DB 1; Length 25;
Best Local Similarity 66.7%; Pred. No. 0.077;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFOLKKNMKVYR 12
   ||| :|||
   ||| :|||
Db 4 CFQWRNMRKVR 15
```



## RESULT 15

US-07-891-174-10  
Sequence 10, Application US/07891174  
Patent No. 5317084  
GENERAL INFORMATION:  
APPLICANT: Mamoru TOMITA et al.  
TITLE OF INVENTION: Antimicrobial Peptides and an  
TITLE OF INVENTION: Antimicrobial Agent  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500Kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/891,174  
FILING DATE: 29-MAY-1992  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/755,161  
FILING DATE: 05-SEP-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX: 202-371-8856  
TELEX:  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM:  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 4  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 4 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 21"  
FEATURE:  
NAME/KEY: modified site  
LOCATION: 21

## IDENTIFICATION METHOD:

OTHER INFORMATION: /note= "thiol group of  
OTHER INFORMATION: Cys residue at location 21 connected by disulfide bond with  
OTHER INFORMATION: thiol group of Cys residue at location 4"  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-891-174-10  
Query Match 75.0%; Score 48; DB 1; Length 25;  
Best Local Similarity 66.7%; Pred. NO. 0.077;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CFOLKXNMKKVR 12  
Db 4 CFOWQRMNRKVR 15  
Search completed: February 21, 2003, 08:04:28  
Job time : 9.93023 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 08:08:15 ; Search time 6.89372 Seconds  
(without alignments)  
54.162 Million cell updates/sec

Title: US-09-743-107b-96

Perfect score: 64

Sequence: 1 CFOLKKQWKVKR 12

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 156504 seqs, 31069816 residues

Total number of hits satisfying chosen parameters: 156504

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:\*
- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW PUB.pep.\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW PUB.pep.\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW PUB.pep.\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW PUB.pep.\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW PUB.pep.\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW PUB.pep.\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW PUB.pep.\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	75.0	15	9	US-09-798-869-2
2	48	75.0	25	9	US-09-798-869-20
3	48	75.0	694	9	US-10-023-096-2
4	43	67.2	15	9	US-09-798-869-6
5	42	65.6	351	10	US-09-853-625B-16
6	39	60.9	57	10	US-09-864-761-37163
7	37	57.8	28	9	US-09-974-879-345
8	37	57.8	784	10	US-09-905-983-52
9	36	56.2	358	10	US-09-841-132-539
10	36	56.2	358	12	US-10-007-693-107
11	35	54.7	46	10	US-09-864-761-43420
12	35	54.7	373	10	US-09-760-354A-2
13	34	53.1	347	10	US-09-851-588-4
14	34	53.1	371	9	US-10-046-935-2235
15	34	53.1	371	9	US-09-878-178-2235
16	34	53.1	391	9	US-10-046-935-2239
17	34	53.1	407	10	US-09-925-301-1165
18	33	51.6	35	10	US-09-764-877-1302
19	33	51.6	48	10	US-09-864-761-40172

20	33	51.6	48	10	US-09-071-838-202
21	33	51.6	54	10	US-09-764-846-244
22	33	51.6	65	10	US-09-864-761-42127
23	33	51.6	116	10	US-09-764-853-738
24	33	51.6	193	9	US-09-796-692-2458
25	33	51.6	329	10	US-09-927-738-11
26	33	51.6	333	10	US-09-765-272-58
27	33	51.6	336	9	US-09-987-107-33
28	33	51.6	336	10	US-09-800-729-207
29	33	51.6	401	9	US-09-987-107-36
30	33	51.6	429	9	US-09-987-107-34
31	33	51.6	470	10	US-09-861-451A-14
32	33	51.6	1499	10	US-09-911-836A-2
33	32	50.0	75	9	US-09-981-876-172
34	32	50.0	75	9	US-09-148-543-172
35	32	50.0	152	9	US-09-738-626-5856
36	32	50.0	184	10	US-09-925-301-1248
37	32	50.0	243	10	US-09-815-242-10838
38	32	50.0	347	10	US-09-815-242-11530
39	32	50.0	391	9	US-09-987-107-38
40	32	50.0	391	10	US-09-800-729-208
41	32	50.0	395	9	US-09-987-107-35
42	32	50.0	447	9	US-10-155-789-4
43	32	50.0	447	9	US-10-155-789-17
44	32	50.0	447	12	US-10-042-417-10
45	32	50.0	448	10	US-09-866-562-37

ALIGNMENTS

RESULT 1

US-09-798-869-2  
; Sequence 2, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON  
; APPLICANT: LARS VORLAND  
; TITLE OF INVENTION: BIOACTIVE PEPTIDES  
; FILE REFERENCE: A34049-PCT-USA-A  
; CURRENT APPLICATION NUMBER: US/09/798,869  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/GB99/02851  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: GB9818938.4  
; PRIOR FILING DATE: 1998-08-28  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-09-798-869-2

Query Match 75.0%; Score 48; DB 9; Length 15;  
Best Local Similarity 66.7%; Pred. No. 0.04;  
Matches 8; Conservative 3; Mismatches 1; Indels 0;

Qy 1 CFOLKKQWKVKR 12  
Db 3 CFQWRNRKVR 14

RESULT 2

US-09-798-869-20  
; Sequence 20, Application US/09798869  
; Publication No. US20030022821A1  
; GENERAL INFORMATION:  
; APPLICANT: JOHN SIGURD SVENDSEN  
; APPLICANT: (YSTEIN REKDAL  
; APPLICANT: BALDUR SVEINBJ (RNSON

APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 20  
LENGTH: 25  
TYPE: PRT  
ORGANISM: Homo Sapiens  
US-09-798-869-20

Query Match 75.0%; Score 48; DB 9; Length 25;  
Best Local Similarity 66.7%; Pred. No. 0.065;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 CFOLKKNKKVR 12  
DB 3 CFQWQNRKVR 14

## RESULT 3

US-10-023-096-2  
Sequence 2, Application US/10023096  
Patent No. US20020160941A1  
GENERAL INFORMATION:  
APPLICANT: Kruzel, Marian L.  
APPLICANT: Kurecki, Tomasz  
APPLICANT: Gollnick, Paul D.  
APPLICANT: Doyle, Darrell J.  
TITLE OF INVENTION: Cloning, Expression, and Uses of Human  
TITLE OF INVENTION: Lactoferrin  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jacobson, Price, Holman & Stern  
STREET: 400 Seventh St. N.W.  
CITY: Washington D.C.  
COUNTRY: U.S.A.  
ZIP: 20004

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/023,096  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,586  
FILING DATE: 30-SEPT-1996  
APPLICATION NUMBER: US 08/238,445  
FILING DATE: 05-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Player, William E.  
REGISTRATION NUMBER: 31,409  
REFERENCE/DOCKET NUMBER: 10505/P58185C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 638-6666  
TELEFAX: (202) 393-5350  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 694 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-10-023-096-2

Query Match 75.0%; Score 48; DB 9; Length 694;  
Best Local Similarity 66.7%; Pred. No. 1.5;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 CFOLKKNKKVR 12  
DB 22 CFQWQNRKVR 33

## RESULT 4

US-09-798-869-6  
Sequence 6, Application US/09798869  
Publication No. US2003002821A1  
GENERAL INFORMATION:  
APPLICANT: JOHN SIGURD SVENDSEN  
APPLICANT: (YSTEIN REKDAL  
APPLICANT: BALDUR SVEINEY (RNSSON  
APPLICANT: LARS VORLAND  
TITLE OF INVENTION: BIOACTIVE PEPTIDES  
FILE REFERENCE: A34049-PCT-USA-A  
CURRENT APPLICATION NUMBER: US/09/798,869  
CURRENT FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: PCT/GB99/02851  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: GB9818938.4  
PRIOR FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: SYNTHETIC PEPTIDE (modified form of Homo sapiens  
OTHER INFORMATION: sequence)  
US-09-798-869-6

Query Match 67.2%; Score 43; DB 9; Length 15;  
Best Local Similarity 66.7%; Pred. No. 0.26;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 CFOLKKNKKVR 12  
DB 3 CFQWQNRKVR 14

## RESULT 5

US-09-853-625B-16  
Sequence 16, Application US/09853625B  
Patent No. US20020049304A1  
GENERAL INFORMATION:  
APPLICANT: Hastings, Gregg A. and Adams, Mark D.  
TITLE OF INVENTION: Human CCN-Like Growth Factor  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,  
CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/853,625B  
FILING DATE: 14-May-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/053,587

FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: MULLINS, J.G.  
REGISTRATION NUMBER: 33,073  
REFERENCE/DOCKET NUMBER: 325800-442  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 351 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS: <Unknown>  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
SEQUENCE DESCRIPTION: SEQ ID NO: 16:  
S-09-853-625B-16

Query Match 65.6%; Score 42; DB 10; Length 351;  
Best Local Similarity 66.7%; Pred. No. 7.7;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Y 1 CFOLKKNMKVR 12  
|||:|:|  
b 258 CIQTKSKMKVR 269

RESULT 6  
US-09-864-761-37163  
Sequence 37163, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aconica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 37163  
LENGTH: 57  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL022169.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 9.2  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7.2  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 9.4  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 8  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 10  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 10  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7.5  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.7  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 9.1  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 7.7  
OTHER INFORMATION: EST HUMAN HIT: AW974650.1, EVALUATE 1.00e-26  
OTHER INFORMATION: SWISSPROT HIT: O07920, EVALUATE 2.60e+00  
US-09-864-761-37163

Query Match 60.9%; Score 39; DB 10; Length 57;  
Best Local Similarity 60.0%; Pred. No. 4.2;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFOLKKNMKK 10  
|||:|:|  
Db 47 CFSLSKNLKR 56

RESULT 7  
US-09-974-879-345  
Sequence 345, Application US/09974879  
Publication No. US20030028003A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 125 Human Secreted Proteins  
FILE REFERENCE: P20202  
CURRENT APPLICATION NUMBER: US/09/974,879  
CURRENT FILING DATE: 2001-10-12  
PRIOR APPLICATION NUMBER: US 60/239,893  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: US 09/818,683  
PRIOR FILING DATE: 2001-03-28  
PRIOR APPLICATION NUMBER: US 09/305,736  
PRIOR FILING DATE: 1999-05-05  
PRIOR APPLICATION NUMBER: PCT/US98/23435  
PRIOR FILING DATE: 1998-11-04  
PRIOR APPLICATION NUMBER: US 60/064,911  
PRIOR FILING DATE: 1997-11-07  
PRIOR APPLICATION NUMBER: US 60/064,912  
PRIOR FILING DATE: 1997-11-07  
PRIOR APPLICATION NUMBER: US 60/064,983  
PRIOR FILING DATE: 1997-11-07  
PRIOR APPLICATION NUMBER: US 60/064,900  
PRIOR FILING DATE: 1997-11-07  
PRIOR APPLICATION NUMBER: US 60/064,986  
PRIOR FILING DATE: 1997-11-07  
PRIOR APPLICATION NUMBER: US 60/064,987  
PRIOR FILING DATE: 1997-11-07  
PRIOR APPLICATION NUMBER: US 60/064,908  
PRIOR FILING DATE: 1997-11-07  
PRIOR APPLICATION NUMBER: US 60/064,984  
PRIOR FILING DATE: 1997-11-07  
PRIOR APPLICATION NUMBER: US 60/064,985  
PRIOR FILING DATE: 1997-11-07  
PRIOR APPLICATION NUMBER: US 60/066,094  
PRIOR FILING DATE: 1997-11-17

```

; PRIOR APPLICATION NUMBER: US 60/066,100
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,089
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,095
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US 60/066,090
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 611
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 345
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-974-879-345

Query Match      57.8%; Score 37; DB 9; Length 28;
Best Local Similarity 80.0%; Pred. No. 4.6;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 QLKKNKKVR 12
| | | | |
DB 10 QLKKNKKER 19

RESULT 8
US-09-905-983-52
; Sequence 52, Application US/09905983
; Patent No. US20020045591A1
; GENERAL INFORMATION:
; APPLICANT: Gelger, Benjamin
; APPLICANT: Ben-Ze'ev, Avri
; APPLICANT: Sadot, Einat
; TITLE OF INVENTION: METHODS AND THERAPEUTIC COMPOSITIONS FOR TREATING CANCER
; FILE REFERENCE: 01/22326
; CURRENT APPLICATION NUMBER: US/09/905,983
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-905-983-52

Query Match      57.8%; Score 37; DB 10; Length 784;
Best Local Similarity 77.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 FOLKKNMK 10
| | | | |
DB 379 FOLKKNQKK 387

RESULT 9
US-09-841-132-599
; Sequence 599, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.465C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 599
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis serovar D

US-09-841-132-599
Query Match      56.2%; Score 36; DB 10; Length 358;
Best Local Similarity 60.0%; Pred. No. 75;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFOLKKNMK 10
| | | | |
DB 251 CYVLSKNWEX 260

RESULT 10
US-10-007-693-107
; Sequence 107, Application US/10007693
; Patent No. US20020146766A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; FILE REFERENCE: 210121.515C2
; CURRENT APPLICATION NUMBER: US/10/007,693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 107
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis serovar D
US-10-007-693-107

Query Match      56.2%; Score 36; DB 12; Length 358;
Best Local Similarity 60.0%; Pred. No. 75;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFOLKKNMK 10
| | | | |
DB 251 CYVLSKNWEX 260

RESULT 11
US-09-864-761-43420
; Sequence 43420, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
```

PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 43420  
LENGTH: 46  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC004962.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.94  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.95  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1  
OTHER INFORMATION: SWISSPROT HIT: P05458, EVALUATION 4.80e+00

3-09-864-761-43420  
Query Match 54.7%; Score 35; DB 10; Length 46;  
Best Local Similarity 63.6%; Pred. No. 16;  
Matches 7; Conservative 1; Mismatches 3; Indels 3; Gaps 0;

1 CFQLKKNMKV 11  
|||  
5 CFSKKLMKEV 15

RESULT 12  
3-09-760-354A-2  
Sequence 2, Application US/09760354A  
Patent No. US20020042385A1  
GENERAL INFORMATION:  
APPLICANT: Bergsma, Derek S.  
APPLICANT: Shabon, Usman  
APPLICANT: Elshourbasy, Nabil  
TITLE OF INVENTION: CLONING OF A NOVEL 7TM RECEPTOR AXOR-2  
FILE REFERENCE: GP-70433-C1  
CURRENT APPLICATION NUMBER: US/09/760,354A  
CURRENT FILING DATE: 2001-01-12  
PRIOR APPLICATION NUMBER: US 60/083,034  
PRIOR FILING DATE: 1998-04-24  
PRIOR APPLICATION NUMBER: US 09/277,398  
PRIOR FILING DATE: 1999-03-26  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 373  
TYPE: PRT  
ORGANISM: HOMO SAPIENS

3-09-760-354A-2

Query Match 54.7%; Score 35; DB 10; Length 373;  
Best Local Similarity 60.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

1 CFQLKKNMKV 10  
|||  
340 CFLNKLKK 349

## RESULT 13

US-09-851-588-4  
; Sequence 4, Application US/09851588  
; Patent No. US20020042067A1  
; GENERAL INFORMATION:  
; APPLICANT: Mack, David  
; APPLICANT: Gish, Kurt C.  
; APPLICANT: Wilson, Keith E.  
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND  
; FILE REFERENCE: A-68829-1/DJB/JJD/AMS  
; CURRENT APPLICATION NUMBER: US/09/851,588  
; CURRENT FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: US 09/642,252  
; PRIOR FILING DATE: 2000-08-17  
; PRIOR APPLICATION NUMBER: US 09/656,002  
; PRIOR FILING DATE: 2000-09-06  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 347  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-851-588-4

Query Match 53.1%; Score 34; DB 10; Length 347;  
Best Local Similarity 60.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 QLKKNMKV 12  
::|||  
DB 12 RVKKNLKK 21

## RESULT 14

US-10-046-935-2235  
; Sequence 2235, Application US/10046935  
; Patent No. US20020156011A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yudi  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Wang, Aijun  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.527C1  
; CURRENT APPLICATION NUMBER: US/10/046,935  
; CURRENT FILING DATE: 2002-01-15  
; NUMBER OF SEQ ID NOS: 2239  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2235  
; LENGTH: 371  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: 197, 201, 203, 204  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-10-046-935-2235

Query Match 53.1%; Score 34; DB 9; Length 371;  
Best Local Similarity 60.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 QLKKNMKV 12  
::|||  
DB 12 RVKKNLKK 21

## RESULT 15

```

US-09-878-178-2235
; Sequence 2235, Application US/09878178
; Patent No US20020177552A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527
; CURRENT APPLICATION NUMBER: US/09/878,178
; NUMBER OF SEQ ID NOS: 2237
; CURRENT FILING DATE: 2001-06-08
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2235
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(371)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-878-178-2235

```

```

Query Match      53.1%; Score 34; DB 9; Length 371;
Best Local Similarity 60.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 3 QLKQNMKKVR 12
Db 12 RVKQNLKFR 21

```

```

Search completed: February 21, 2003, 08:11:59
Job time : 6.88372 secs

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.  
M protein - protein search, using sw model  
un on: February 21, 2003, 07:48:01 ; Search time 10.6047 Seconds  
(without alignments)  
108.784 Million cell updates/sec

file: US-09-743-107B-96  
effect score: 64  
sequence: 1 CFQLKKNKKVR 12  
coring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

sarched: 283224 seqs, 96134422 residues  
otal number of hits satisfying chosen parameters: 283224

inimum DB seq length: 0  
aximum DB seq length: 2000000000  
ost-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

atabase : PIR 73:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Match	Length	DB	ID	Description
1	48	75.0	711	1	TFHUL	lactotransferrin p
2	42	65.6	351	2	S20078	NOV protein - chic
3	42	65.6	795	2	T34468	hypothetical prote
4	40	62.5	223	2	T37974	probable peroxisom
5	40	62.5	238	2	T40568	hypothetical prote
6	40	62.5	333	2	T25242	hypothetical prote
7	39	60.9	222	2	T48456	rna binding protei
8	39	60.9	282	2	A97214	uncharacterized co
9	39	60.9	339	2	D90106	hypothetical prote
10	39	60.9	409	2	S47440	secy protein - red
11	39	60.9	1162	2	T21557	hypothetical prote
12	39	60.9	1553	2	T18502	hypothetical prote
13	38	59.4	1166	2	A39432	ATP-dependent deox
14	38	59.4	2014	2	T21560	hypothetical prote
15	38	59.4	2021	2	A84771	similar to ch-TOG
16	37	57.8	22	2	S59071	secretory protein
17	37	57.8	224	2	C31201	GLI-related finger
18	37	57.8	267	2	S77802	hypothetical prote
19	37	57.8	335	2	T33211	hypothetical prote
20	37	57.8	510	2	G70162	conserved hypothet
21	37	57.8	546	2	A99600	conserved hypothet
22	37	57.8	718	2	T29708	hypothetical prote
23	37	57.8	784	1	I0HUC5	cadherin 5 precurs
24	36	56.2	81	2	E85172	hypothetical prote
25	36	56.2	81	2	E88038	hypothetical prote
26	36	56.2	104	2	D95003	hypothetical prote
27	36	56.2	191	2	B81220	cell filamentation
28	36	56.2	227	2	A59097	hypothetical prote
29	36	56.2	258	2	T00447	hypothetical prote

30	36	56.2	260	2	T00446	hypothetical prote
31	36	56.2	282	2	S50454	hypothetical prote
32	36	56.2	294	2	S49781	hypothetical prote
33	36	56.2	295	2	D86255	hypothetical prote
34	36	56.2	315	2	AD1143	B. subtilis Deor t
35	36	56.2	358	2	D71533	probable p-loop AT
36	36	56.2	358	2	D71533	hypothetical prote
37	36	56.2	551	2	T22121	hypothetical prote
38	36	56.2	566	1	F84329	arginine-tRNA liga
39	36	56.2	667	2	T15710	hypothetical prote
40	36	56.2	887	2	A96516	hypothetical prote
41	36	56.2	991	2	A99334	hypothetical prote
42	36	56.2	1086	2	T18294	Ca2+-transporting
43	35	54.7	68	2	F90520	hypothetical prote
44	35	54.7	99	2	T22865	hypothetical prote
45	35	54.7	107	2	G90017	hypothetical prote
	35	54.7	145	2	C90107	hypothetical prote

ALIGNMENTS

RESULT 1

TFHUL  
lactotransferrin precursor [validated] - human  
N/Alternate names: lactoferrin  
C/Species: Homo sapiens (man)  
C/Date: 31-Mar-1992 #sequence revision 21-Nov-1997 #text change 08-Dec-2000  
C/Accession: G01394; S11228; A45401; S10324; S15853; S20841; S07160; A61189; A31000; S741  
R/Cho, Y.  
submitted to the EMBL Data Library, March 1994  
A/Reference number: G06820  
A/Accession: G01394  
A/Status: Preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-711 <CHO>  
A/Cross-references: EMBL:U07643; NID:G467236; PIDN:AAB60324.1; PID:G467237  
R/Key, M.W.; Woloshuk, S.L.; deBoer, H.A.; Pieper, F.R.  
Nucleic Acids Res. 18, 5288, 1990  
A/Title: Complete nucleotide sequence of human mammary gland lactoferrin.  
A/Reference number: S11228; MUID:90384839; PMID:2402455  
A/Accession: S11228  
A/Molecule type: mRNA  
A/Residues: 1-48, 'T', 150-422, 'C', 424-711 <REY>  
A/Cross-references: EMBL:X53961; NID:G34415; PIDN:CAA37914.1; PID:G34416  
R/Teng, C.T.; Liu, Y.; Yang, N.; Walmer, D.; Panella, T.  
Mol. Endocrinol. 6, 1969-1981, 1992  
A/Title: Differential molecular mechanism of the estrogen action that regulates lactofer:  
A/Reference number: A45401; MUID:93125571; PMID:1480183  
A/Accession: A45401  
A/Molecule type: DNA  
A/Residues: 1-15 <PEN>  
A/Cross-references: GB:S52659; NID:G263311; PIDN:AAB24877.1; PID:G263312  
A/Experimental source: Placenta  
A/Note: sequence extracted from NCBI backbone (NCBIFP:122202)  
R/Powell, M.J.; Ogden, J.E.  
Nucleic Acids Res. 18, 4013, 1990  
A/Title: Nucleotide sequence of human lactoferrin cDNA.  
A/Reference number: S10324; MUID:90326549; PMID:2374734  
A/Accession: S10324  
A/Molecule type: mRNA  
A/Residues: 3-711 <POW>  
A/Cross-references: EMBL:X52941; NID:G34411; PIDN:CAA37116.1; PID:G34412  
R/Stowell, K.M.; Rado, T.A.; Funk, W.D.; Tweedie, J.W.  
Biochem. J. 276, 349-355, 1991  
A/Title: Expression of cloned human lactoferrin in baby-hamster kidney cells.  
A/Reference number: S15853; MUID:91264786; PMID:2049066  
A/Accession: S15853  
A/Status: nucleic acid sequence not shown; not compared with conceptual translation  
A/Molecule type: mRNA  
A/Residues: 20-31 <STI>  
A/Accession: S20841  
A/Molecule type: protein  
A/Residues: 20-28, 'X', 30-31 <ST2>



```

Query Match          65.6%; Score 42; DB 2; Length 351;
Best Local Similarity 66.7%; Pred. No. 9.5;
Matches      8; Conservative      1; Mismatches      3; Indels      0; Gaps      0;

QY      1 CFOLKKNMKVR 12
      |||:|:|
Db      258 CIQTQKSMKAVR 269

RESULT 3
T34468
hypothetical protein ZK770.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T34468
R:Maggi, L.; Gattung, S.; Bartko, L.
A:Submitted to the EMBL Data Library, April 1997
A:Description: The sequence of C. elegans cosmid ZK770.
A:Reference number: Z21530
A:Accession: T34468
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-795 <MAG>
A:Cross-references: EMBL:U97404; PIDN:AAB93303.1; GSPDB:GNO0019; CESP:ZK770.1
A:Experimental source: strain Bristol N2; clone ZK770
C:Genetics:
A:Gene: CESP:ZK770.1
A:Map position: 1
A:Introns: 18/1, 84/2, 125/3, 198/3, 257/3, 337/3, 378/3, 424/3, 477/2, 510/1, 563/

Query Match          65.6%; Score 42; DB 2; Length 795;
Best Local Similarity 70.0%; Pred. No. 20;
Matches      7; Conservative      2; Mismatches      1; Indels      0; Gaps      0;

QY      1 CFOLKKNMKK 10
      |||:|:|
Db      329 CFQIKGNVK 338

RESULT 4
T37974
probable peroxisomal membrane protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000
C:Accession: T37974
R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
A:Submitted to the EMBL Data Library, March 1996
A:Reference number: Z21759
A:Accession: T37974
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-223 <MUR>
A:Cross-references: EMBL:Z69909; PIDN:CAA93785.1; GSPDB:GNO0066; SPDB:SPAC19G10.03c
A:Experimental source: strain 972h-; cosmid cl9G10
C:Genetics:
A:Gene: SPDB:SPAC19G10.03c
A:Map position: 1
A:Introns: 10/3, 170/2
C:Superfamily: Schizosaccharomyces pombe hypothetical protein SPBC582.09
C:Keywords: peroxisome

Query Match          62.5%; Score 40; DB 2; Length 223;
Best Local Similarity 41.7%; Pred. No. 14;
Matches      5; Conservative      6; Mismatches      1; Indels      0; Gaps      0;

QY      1 CFOLKKNMKVR 12
      ||:|:|:|
Db      140 CYELOQNSKKIK 151

RESULT 5
T40568

```

A;Cross-references: EMBL:AL162875  
A;Experimental source: cultivar Columbia; BAC clone T32M21  
C;Genetics:  
A;Map position: 5  
A;Introns: 38/3; 78/1; 96/3; 112/3; 145/2; 168/3; 205/3  
A;Note: T32M21.200

Query Match 60.9%; Score 39; DB 2; Length 222;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LKKNMKKV 11  
| | | | |  
DB 9 LKKNMKKV 16  
| | | | |

RESULT 8  
A97214  
uncharacterized conserved protein CAC2549 [imported] - Clostridium acetobutylicum  
C;Species: Clostridium acetobutylicum  
C;Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 30-Sep-2001  
C;Accession: A97214  
R;Nolling, J.; Brenner, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum  
A;Reference number: A96900; MUID:21359325; PMID:21359325  
A;Accession: A97214  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-282 <KUR>  
A;Cross-references: GB:AE001437; PIDN:AAK80500.1; PID:g15025572; GSPDB:GN00168  
A;Experimental source: Clostridium acetobutylicum ATCC824  
C;Genetics:  
A;Gene: CAC2549  
C;Superfamily: Deinococcus radiodurans hypothetical protein DRB00099

Query Match 60.9%; Score 39; DB 2; Length 282;  
Best Local Similarity 58.3%; Pred. No. 26;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFOLKKNMKKV 12  
| | | | | | | |  
DB 241 CQOLKENISKLR 252  
| | | | | | | |

RESULT 9  
D90106  
hypothetical protein orf339 [imported] - Guillardia theta nucleomorph  
C;Species: nucleomorph Guillardia theta  
A;Date: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
C;Note: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001  
C;Accession: D90106  
R;Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reiter, A.; The highly reduced genome of an enslaved algal nucleus.  
Nature 410, 1091-1096, 2001  
A;Reference number: A99082; MUID:l1323671; PMID:l1323671  
A;Accession: D90106  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-339 <DOU>  
A;Cross-references: GB:AJ010592; NID:g12580683; PIDN:CAC27001.1; GSPDB:GN00151  
C;Genetics:  
A;Gene: orf339  
A;Map position: 2  
A;Genome: nucleomorph  
C;Keywords: nucleomorph

Query Match 60.9%; Score 39; DB 2; Length 339;  
Best Local Similarity 80.0%; Pred. No. 31;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 FOLKKNMKKV 11

```
Db 179 FNLKKNKKV 188
|||||
Query Match 60.9%; Score 39; DB 2; Length 409;
Best Local Similarity 60.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

RESULT 10
S47440
SecY protein - red alga (Cyanidium caldarium)
C:Species: Cyanidium caldarium
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Sep-1999
C:Accession: S47440
R:Vogel, H.; Valentini, K.
submitted to the EMBL Data Library, August 1994
A:Reference number: S47440
A:Accession: S47440
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-409 <VOG>
A:Cross-references: EMBL:Z36235; NID:G529651; PIDN:CAA85270.1; PID:G529652
C:Superfamily: preprotein translocase secY

Query Match 60.9%; Score 39; DB 2; Length 409;
Best Local Similarity 60.0%; Pred. No. 36;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFOLKKNKK 10
|||||
Db 312 CPFLSNLKK 321
|||||

RESULT 11
T21557
Hypothetical protein F29G6.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T21557
R:Haris, B.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z19441
A:Accession: T21557
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1162 <WIL>
A:Cross-references: EMBL:Z78543; PIDN:CAB01753.1; GSPDB:GN00028; CBSP:F29G6.1
A:Experimental source: clone F29G6
C:Genetics:
A:Gene: CBSP:F29G6.1
A:Map position: X
A:Introns: 51/1; 78/1; 132/2; 161/1; 236/3; 276/3; 318/3; 359/3; 397/3; 470/3; 544/3

Query Match 60.9%; Score 39; DB 2; Length 1162;
Best Local Similarity 50.0%; Pred. No. 93;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFOLKKNKKV 12
|||||
Db 943 CIQMKKNKTIQ 954
|||||

RESULT 12
T18502
Hypothetical protein C0755c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C:Accession: T18502
R:Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z18935
A:Accession: T18502
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1553 <LAW>
A:Cross-references: EMBL:Z98551; NID:G1331903; PID:el331923; PIDN:CAB11141.1
C:Genetics:

us-09-743-107b-96.rpr
A:Map position: 3
A>Note: C0755c

Query Match 50.9%; Score 39; DB 2; Length 1553;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 LKKNKKV 12
|||||
Db 39 MKKNKKV 47
|||||

RESULT 13
A39432
ATP-dependent deoxyribonuclease chain B - Bacillus subtilis
N:Alternate names: ATP-dependent exonuclease synthesis protein AddB
C:Species: Bacillus subtilis
C>Date: 21-Feb-1992 #sequence_revision 21-Feb-1992 #text_change 02-Feb-2001
C:Accession: A39432; S61272; A69583
R:Koolstra, J.; Venema, G.
J. Bacteriol. 173, 3644-3655, 1991
A:Title: Cloning, sequencing, and expression of Bacillus subtilis genes involved in ATP-
A:Reference number: A39432; MUID:91267926; PMID:1646786
A:Accession: A39432
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1166 <KOO>
A:Cross-references: GB:M63489; NID:G142438; PIDN:AAA22200.1; PID:G142439
R:Hajjema, B.J.; Hamoen, L.W.; Koolstra, J.; Venema, G.; van Sinderen, D.
Mol. Microbiol. 15, 203-211, 1995
A:Title: Expression of the ATP-dependent deoxyribonuclease of Bacillus subtilis is under
A:Reference number: S61272; MUID:95264907; PMID:7746142
A:Accession: S61272
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-13 <HAI>
R:Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C:Bron, S.; Brouillet, S.; Bruschini, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A:Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Puma, S.; Galizzi, A.; Galler
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Hoisappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maguda, S.; Maueel
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: A69583
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1166 <KUN>
A:Cross-references: GB:Z99109; GB:AL009126; NID:G2633260; PIDN:CAB12902.1; PID:G2633398
A:Experimental source: strain 168
C:Genetics:
A:Gene: addB
C:Superfamily: Bacillus subtilis ATP-dependent deoxyribonuclease chain B
C:Keywords: nucleotide binding; P-loop
F:8-15/Region: nucleotide-binding motif A (P-loop)

Query Match 59.4%; Score 38; DB 2; Length 1166;
Best Local Similarity 63.6%; Pred. No. 1.4e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 FOLKKNKKV 12
|||||
Db 486 FLOKKNKKK 496
|||||
```

```

LT 14
360
>ethical protein T25C12.3 - Caenorhabditis elegans
>ecies: Caenorhabditis elegans
>te: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
>cession: T21560; T23053; T25270
>ris, B.
>itted to the EMBL Data Library, August 1996
>ference number: Z19441
>cession: T21560
>atus: preliminary; translated from GB/EMBL/DBJ
>ecule type: DNA
>idues: 1-2014 <WIL>
>ss-references: EMBL:Z78543; PIDN:CAB01757.1; GSPDB:GN000028; CESP:T25C12.3
>erimental source: clone F29G6
>ite, S.
>itted to the EMBL Data Library, April 1997
>ference number: Z19664
>cession: T23053
>atus: preliminary; translated from GB/EMBL/DBJ
>ecule type: DNA
>idues: 1-2014 <W12>
>ss-references: EMBL:Z33779; PIDN:CAB07849.1; GSPDB:GN000028; CESP:T25C12.3
>erimental source: clone H06X08
>shaw, J.
>itted to the EMBL Data Library, November 1995
>ference number: Z20007
>cession: T25270
>atus: preliminary; translated from GB/EMBL/DBJ
>ecule type: DNA
>idues: 1-2014 <W13>
>ss-references: EMBL:Z66566; PIDN:CAA91487.1; GSPDB:GN000028; CESP:T25C12.3
>erimental source: clone T25C12
>netics:
>ne: CESP:T25C12.3
>p position: X
>trons: 10/2; 91/3; 136/1; 159/2; 182/3; 203/1; 225/3; 280/1; 422/3; 466/3; 559/3; 69
>22/1; 1779/3; 1843/3; 1908/2; 1941/2; 2007/2

ery Match 59.4%; Score 38; DB 2; Length 2014;
>st Local Similarity 87.5%; Pred. No. 2.3e+02;
>tches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

2 FOLKKNMK 9
|||||:|
1410 FOLKKNLK 1417

LT 15
71
lar to ch-TOG protein from Homo sapiens [imported] - Arabidopsis thaliana
>ecies: Arabidopsis thaliana (mouse-ear cress)
>te: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
>cession: A84771
>n, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
>Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.;
>D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
>re 402, 761-768, 1999
>tle: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
>ference number: A84420; MUID:20083487; PMID:10617197
>cession: A84771
>atus: preliminary
>ecule type: DNA
>idues: 1-2021 <STO>
>ss-references: GB:AE002093; NID:G4263790; PIDN:ADL5450.1; GSPDB:GN00139
>netics:
>ne: At2g35630
>p position: 2

ery Match 59.4%; Score 38; DB 2; Length 2021;
>st Local Similarity 58.3%; Pred. No. 2.3e+02;
>tches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

Qy 1 CFOLKKNMKKVR 12  
 |||||:  
 Db 1405 CFRLAKDMKKRR 1416

Search completed: February 21, 2003, 08:02:51  
 Job time : 12.6047 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

M protein - protein search, using sw model

un on: February 21, 2003, 07:28:06 ; Search time 5.2093 Seconds  
(without alignments)  
95.544 Million cell updates/sec

itle: US-09-743-107B-96  
effect score: 64  
sequence: 1 CFQLKKNMKVR 12

coring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

earched: 112892 seqs, 41476328 residues

otal number of hits satisfying chosen parameters: 112892

inimum DB seq length: 0  
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

atabase : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

result No.	Score	Match	Length	DB ID	Description
1	48	75.0	711	1	TRFL_HUMAN
2	42	65.6	351	1	P26788 homo sapien
3	42	65.6	573	1	P26886 gallus gall
4	42	65.6	795	1	P26929 lactobacilli
5	41	64.1	695	1	O01635 caenorhabdi
6	40	62.5	238	1	O77811 equus caball
7	39	60.9	410	1	Q10333 schizosacch
8	38	59.4	353	1	P46249 cyanidium c
9	38	59.4	1166	1	P32642 coturnix co
10	37	57.8	267	1	P23477 bacillus su
11	37	57.8	510	1	P3661 mycoplasma
12	37	57.8	697	1	O51457 borrelia bu
13	37	57.8	708	1	P10072 homo sapien
14	37	57.8	784	1	Q8tum0 camelus dro
15	36	56.2	282	1	P33151 homo sapien
16	36	56.2	258	1	P40003 saccharomyc
17	36	56.2	560	1	O84289 chlamydia t
18	36	56.2	566	1	P04665 influenza b
19	35	54.7	175	1	Q57689 methanococc
20	35	54.7	195	1	O51372 borrelia bu
21	35	54.7	352	1	O83049 treponema p
22	35	54.7	373	1	Q9n66 homo sapien
23	35	54.7	373	1	Q5j112 rattus norv
24	35	54.7	414	1	Q81122 methanococc
25	35	54.7	656	1	P51392 porphyra pu
26	35	54.7	1008	1	P36048 saccharomyc
27	35	54.7	4967	1	RYR2_HUMAN
28	34	53.1	126	1	RK12_CVAPA
29	34	53.1	174	1	FTRV_SPIOL
30	34	53.1	181	1	TRIF_HUMAN
31	34	53.1	181	1	TRIF_MOUSE
32	34	53.1	181	1	TRIF_RABIT
33	34	53.1	181	1	P27768 rattus norv

## ALIGNMENTS

RESULT 1	TRFL_HUMAN	STANDARD;	PRT;	711 AA.
ID	P02788; Q16780; Q16785; Q16789; Q00756; Q9H123; Q96KZ4;			
AC	Q96KZ5;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Lactotransferrin precursor (Lactoferrin) [Contains: Lactoferrin A;			
DE	Lactoferrin B; Lactoferrin C]			
GN	LTF OR LF.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1] RN			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RC	MEDLINE=90384839; PubMed=2402455;			
RA	Rey M.W., Woloshuk S.L., de Boer H.A., Pieper F.R.;			
RT	"Complete nucleotide sequence of human mammary gland lactoferrin.";			
RL	Nucleic Acids Res. 18:5288-5288(1990).			
RN	[2] RN			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Choi Y.Y.;			
RL	Thesis (1994), Genetic Engineering Research Institute / Taejeon, Korea.			
RN	[3] RN			
RP	SEQUENCE FROM N.A.			
RA	Conceely O.M.;			
RL	Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.			
RN	[4] RN			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Liang Q., Jimenez-Flores R., Richardson T.;			
RT	"Molecular cloning and sequence analysis of human lactoferrin.";			
RL	Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.			
RN	[5] RN			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Bone marrow;			
RA	Wei X., Han J., Rado T.A.;			
RT	"Human neutrophil lactoferrin coding and 5' flanking region DNA			
RL	sequences.";			
RN	[6] RN			
RP	Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.			
RN	[6] RN			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RA	Cheng H., Chen X., Huan L.;			
RT	"cDNA cloning and sequence analysis of human lactoferrin.";			
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RN	[7] RN			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Prostate;			
RA	Strausberg R.;			
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.			
RN	[8] RN			

P77577 escherichia  
Q58684 methanococc  
P53210 saccharomyc  
Q91918 brachydanio  
P11102 influenza b  
P13884 influenza b  
P13885 influenza b  
Q36433 influenza b  
P04666 influenza b  
Q16204 homo sapien  
P52376 human herpe  
P12465 chlorella v

34 34 53.1 228 1 YDFH\_ECOLI  
35 34 53.1 252 1 YC88\_METUA  
36 34 53.1 297 1 YG1B\_YEAST  
37 34 53.1 387 1 SRB3\_BRARE  
38 34 53.1 560 1 VNUC\_INBAA  
39 34 53.1 560 1 VNUC\_INBAC  
40 34 53.1 560 1 VNUC\_INBAD  
41 34 53.1 560 1 VNUC\_INBP9  
42 34 53.1 560 1 VNUC\_INBSI  
43 34 53.1 585 1 D170\_HUMAN  
44 34 53.1 659 1 HEPA\_HSV7J  
45 34 53.1 1562 1 RPOD\_CHLVU

RP SEQUENCE OF 3-711 FROM N.A.  
RC TISSUE=mammary gland;  
RX MEDLINE=90326549; PubMed=2374734;  
RA Powell M.J., Ogden J.E.;  
RT "Nucleotide sequence of human lactoferrin cDNA.";  
RN Nucleic Acids Res. 18:4013-4013(1990).  
RP SEQUENCE OF 20-711.  
RX MEDLINE=8507667; PubMed=6510420;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Schoentgen F.,  
RG Legrand D., Spik G., Montreuil J., Jolles P.;  
RT "Human lactoferrin: amino acid sequence and structural  
RT comparisons with other transferrins.";  
RL Eur. J. Biochem. 145:659-666(1984).  
RN [10]  
RP PRELIMINARY SEQUENCE OF 20-73; 134-171; 257-278 AND 347-530.  
RX MEDLINE=82046817; PubMed=6794640;  
RA Metz-Boutigue M.-H., Mazurier J., Jolles J., Spik G., Montreuil J.,  
RG Jolles P.;  
RT "The present state of the human lactotransferrin sequence. Study and  
RT alignment of the cyanogen bromide fragments and characterization of  
RT N- and C-terminal domains.";  
RL Biochim. Biophys. Acta 670:243-254(1981).  
RN [11]  
RP SEQUENCE OF 609-711.  
RX MEDLINE=82262043; PubMed=7049727;  
RA Metz-Boutigue M.-H., Jolles J., Mazurier J., Spik G., Montreuil J.,  
RG Jolles P.;  
RT "An 88 amino acid long C-terminal sequence of human  
RT lactotransferrin.";  
RL FEBS Lett. 142:107-110(1982).  
RN [12]  
RP SEQUENCE OF 436-711 FROM N.A.  
RX MEDLINE=86001031; PubMed=3477300;  
RA Rado T.A., Wei X., Benz E.J. Jr.;  
RT "Isolation of lactoferrin cDNA from a human myeloid library and  
RT expression of mRNA during normal and leukemic myelopoiesis.";  
RL Blood 70:989-993(1987).  
RN [13]  
RP SEQUENCE OF 237-711 FROM N.A.  
RA McCombie W.R., Wilson R., Chen E., Gibbs R., Zuo L., Johnson D.,  
RG Nham M., Parnell L., Dedhia N., Ansari A., Mardis E., Schutz K.,  
RA Gnoj L., la Baside M., Kaplan N., Greco T., Touchman J., Musny D.,  
RG Chen C.N., Evans C., Fitzgerald M., See L.H., Tang M., Porcel B.M.,  
RA Dragan Y., Giacalone J., Pae A., Powell E., Solinsky K.A., Desilva U.,  
RG Diaz-Perez S., Zhou X., Yu Y., Watanabe M., Doggett N., Garcia D.,  
RA Sacripanti J.L.;  
RT Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
RN [14]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS), AND REVISIONS.  
RX MEDLINE=90064528; PubMed=2585506;  
RA Anderson B.F., Baker H.M., Norris G.E., Rice D.W., Baker E.N.;  
RT "Structure of human lactoferrin: crystallographic structure analysis  
RT and refinement at 2.8-A resolution.";  
RL J. Mol. Biol. 209:711-734(1989).  
RN [15]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX MEDLINE=97156796; PubMed=9003186;  
RA Nicholson H., Anderson B.F., Bland T., Shewry S.C., Tweedie J.W.,  
RG Baker E.N.;  
RT "Mutagenesis of the histidine ligand in human lactoferrin: iron  
RT binding properties and crystal structure of the histidine-  
RT 253->methionine mutant.";  
RL Biochemistry 36:1341-1346(1997).  
RN [17]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX MEDLINE=99190892; PubMed=10089347;  
RA Sun X.L., Baker H.M., Shewry S.C., Jameson G.B., Baker E.N.;  
RT "Structure of recombinant human lactoferrin expressed in Aspergillus  
RT awamori.";  
RL Acta Crystallogr. D 55:403-407(1999).  
RN [18]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
RX MEDLINE=99192677; PubMed=10089508;  
RA Jameson G.B., Anderson B.F., Norris G.E., Thomas D.H., Baker E.N.;  
RT "Structure of human apolactoferrin at 2.0-A resolution. Refinement  
RT and analysis of ligand-induced conformational change.";  
RL Acta Crystallogr. D 54:1319-1335(1998).  
RN [19]  
RP CHARACTERIZATION OF LACTOFERROXINS.  
RX MEDLINE=91166929; PubMed=1369293;  
RA Tani F., Iio K., Chiba H., Yoshikawa M.;  
RT "Isolation and characterization of opioid antagonist peptides derived  
RT from human lactoferrin.";  
RL Agric. Biol. Chem. 54:1803-1810(1990).  
RN [20]  
RP VARIANTS THR-30 AND ARG-48.  
RX PubMed=9873069;  
RA Klintworth G.K., Sommer J.R., Obrian G., Han L., Ahmed M.N.,  
RG Qumsiyeh M.B., Lin P.-Y., Basti S., Reddy M.K., Kanai A., Hotta Y.,  
RA Sugar J., Kumaramanickavel G., Munier F., Schorderet D.F.,  
RG El Matri L., Iwata F., Kaiser-Kupfer M., Nagata M., Nakayasu K.,  
RA Hejtmancik J.F., Teng C.T.;  
RT "Familial subepithelial corneal amyloidosis (gelatinous drop-like  
RT corneal dystrophy): exclusion of linkage to lactoferrin gene.";  
RL Mol. Vision 4:31-32(1998).  
RN [21]  
RP FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
RP CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
RP OF AN ANION, USUALLY BICARBONATE.  
RP FUNCTION: LACTOFERROXIN A, B AND C HAVE OPIOID ANTAGONIST  
RP ACTIVITY. LACTOFERRONIN A SHOWS PREFERENCE FOR MU-RECEPTORS, WHILE  
RP LACTOFERRONIN B AND C HAVE SOMEWHAT HIGHER DEGREES OF PREFERENCE  
RP FOR KAPPA-RECEPTORS THAN FOR MU-RECEPTORS.  
RP SUBUNIT: MONOMER.  
RP SUBCELLULAR LOCATION: Secreted.  
RP DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
RP SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
RP This SWISS-PROT entry is copyright. It is produced through a collaboration  
RP between the Swiss Institute of Bioinformatics and the EMBL outstation -  
RP the European Bioinformatics Institute. There are no restrictions on its  
RP use by non-profit institutions as long as its content is in no way  
RP modified and this statement is not removed. Usage by and for commercial  
RP entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
RP or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
RN [22]  
RP EMBL; X53961; CAA37914.1; -  
RN EMBL; U07643; AAB60324.1; -  
RN EMBL; M93150; AAA36159.1; -  
RN EMBL; M83202; AAA59511.1; -  
RN EMBL; M83205; AAA58656.1; -  
RN EMBL; M18642; AAA86665.1; -  
RN EMBL; AF332168; AAG48753.1; -  
RN EMBL; BC015822; AAH15822.1; -  
RN EMBL; BC015823; AAH15823.1; -  
RN EMBL; X52941; CAA37116.1; -  
RN EMBL; U95626; AAB57795.1; -  
RN PIR; S11228; TPFUL.  
RN PDB; 1LCF; 31-AUG-94.  
RN PDB; 1LCT; 31-OCT-93.  
RN PDB; 1LFG; 31-JUL-94.  
RN PDB; 1LFH; 31-OCT-93.  
RN PDB; 1LFI; 31-OCT-93.  
RN PDB; 1LGB; 31-AUG-94.  
RN PDB; 1LGC; 31-AUG-94.  
RN PDB; 1BKA; 08-NOV-96.  
RN PDB; 1DSN; 08-MAR-96.  
RN PDB; 1HSE; 12-MAR-97.  
RN PDB; 1VFD; 21-APR-97.



PT METAL 365 365 NICKEL 2 (BY SIMILARITY).  
 QT ACT SITE 325 325 BY SIMILARITY.  
 QZ SEQUENCE 573 AA; 61805 MW; 2D2619781C39E54B CRC64;

Query Match 55.6%; Score 42; DB 1; Length 573;  
 Best Local Similarity 56.7%; Pred. No. 5.5;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Y 1 CFQLKKNMKYR 12  
 |||||  
 b 379 CQWLASKMKYR 390

RESULT 4  
 DEGY CAEEL STANDARD; PRT; 795 AA.  
 IC 001635;  
 VT 15-JUL-1998 (Rel. 36, Created)  
 VT 15-JUL-1998 (Rel. 36, Last sequence update)  
 JT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Degenerin-like protein ZK770.1 in chromosome I.  
 N ZK770.1.  
 S Caenorhabditis elegans.  
 C Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditoidea;  
 C Rhabditidae; Peloderinae; Caenorhabditis.  
 X NCBI\_TaxID=6239;  
 N [1]  
 P SEQUENCE FROM N.A.  
 C STRAIN=Bristol N2;  
 A Maggi L., Gattung S., Bartko L.;  
 L Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 C -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 C -1- SIMILARITY: BELONGS TO THE AMILORIDE-SENSITIVE SODIUM CHANNEL FAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; U97404; AAB93309.1; --  
 WormPep; ZK770.1; CE1541.  
 R InterPro; IPR004725; Deg-1.  
 R InterPro; IPR001873; Na-channel\_ASC.  
 R Pfam; PF00858; ASC; 1.  
 R PRINTS; PR01078; AMINACHANNEL.  
 R TIGRFAMs; TIGR00867; deg-1; 1.  
 R PROSITE; PS01206; ASC; 1.  
 M Hypothetical protein; Ionic channel; Transmembrane; Ion transport; Glycoprotein.  
 M DOMAIN 1 58 CYTOPLASMIC (POTENTIAL).  
 T TRANSMEM 59 79 POTENTIAL.  
 T DOMAIN 80 739 EXTRACELLULAR (POTENTIAL).  
 T TRANSMEM 740 760 POTENTIAL.  
 T DOMAIN 761 795 CYTOPLASMIC (POTENTIAL).  
 T CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 T CARBOHYD 367 367 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 T CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 T CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 T CARBOHYD 499 499 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 T CARBOHYD 518 518 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 Q SEQUENCE 795 AA; 90776 MW; 2CAACF7D41770B54 CRC64;

Query Match 55.6%; Score 42; DB 1; Length 795;  
 Best Local Similarity 70.0%; Pred. No. 7.4;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Y 1 CFQLKKNMKYR 10  
 |||||  
 b 329 CFQIKGNVKK 338

RESULT 5  
 TRFL HORSE STANDARD; PRT; 695 AA.  
 AC 077811;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (Lactoferrin) (Fragment).  
 GN LTF.  
 OS Equus caballus (Horse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OC NCBI\_TaxID=9796;  
 RN [1]  
 P SEQUENCE FROM N.A.  
 R Paramasivam M., Srinivasan A., Yadav M.P., Singh T.P.;  
 RA "CDNA sequence of mare lactoferrin";  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).  
 RC TISSUE=Milk;  
 RX MEDLINE=99296831; PubMed=10366507;  
 RA Sharma A.K., Paramasivam M., Srinivasan A., Yadav M.P., Singh T.P.;  
 RT "Three-dimensional structure of mare diferric lactoferrin at 2.6-A resolution.";  
 RT J. Mol. Biol. 289:303-317(1999).  
 CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING OF AN ANION, USUALLY BICARBONATE.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; AJ010930; CAA09407.1; --  
 DR PDB; 1BIX; 02-DEC-98.  
 DR PDB; 1B7U; 02-FEB-99.  
 DR PDB; 1B7Z; 02-FEB-99.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 2.  
 DR PRINTS; PR00422; TRANSFERRIN.  
 DR SMART; SM00094; TR\_FER; 2.  
 DR PROSITE; PS00205; TRANSFERRIN\_1; 2.  
 DR PROSITE; PS00206; TRANSFERRIN\_2; 2.  
 DR PROSITE; PS00207; TRANSFERRIN\_3; 1.  
 M Signal; Iron transport; Glycoprotein; Metal-binding; Repeat; Signal; 3D-structure.  
 M NON\_TER 1 6  
 M SIGNAL <1 6  
 M CHAIN 7 695 LACTOTRANSFERRIN.  
 M REPEAT 7 350 1.  
 M REPEAT 351 695 2.  
 M FT DISULFID 15 51  
 M FT DISULFID 25 42  
 M FT DISULFID 121 204  
 M FT DISULFID 163 179  
 M FT DISULFID 186 189  
 M FT DISULFID 176 187  
 M FT DISULFID 237 251  
 M FT DISULFID 354 386  
 M FT DISULFID 364 377  
 M FT DISULFID 411 690  
 M FT DISULFID 431 653



**T** DISULFID 463  
**T** DISULFID 487  
**T** DISULFID 497  
**T** DISULFID 508  
**T** DISULFID 579  
**T** DISULFID 631  
**T** METAL 66  
**T** METAL 98  
**T** METAL 198  
**T** METAL 259  
**T** METAL 401  
**T** METAL 439  
**T** METAL 532  
**T** METAL 601  
**T** BINDING 127  
**T** BINDING 469  
**T** CARBOHYD 143  
**T** CARBOHYD 287  
**T** CARBOHYD 482  
**C** SEQUENCE 695 AA; 75991 MW; 07BB84D50E1B165D CRC64;

Query Match 64.1%; Score 41; DB 1; Length 695;  
Best Local Similarity 58.3%; Pred. No. 9.8;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

**Y** 1 CFQLKKMKVKVR 12  
| : : | | | |  
**D** 25 CAKFQNMKVKVR 36

**RESULT 6**  
**RM9-SCHPO** STANDARD; PRT; 238 AA.  
**O10333**; Rel. 34, Created)  
16-OCT-1996 (Rel. 40, Last sequence update)  
15-JUN-2002 (Rel. 41, Last annotation update)  
Hypothetical protein C582.09 in chromosome II.  
SPBC582.09.  
Schizosaccharomyces pombe (Pission yeast).  
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
Schizosaccharomycetaceae; Schizosaccharomycetaceae;  
Schizosaccharomycetes.  
NCBI\_TaxID=4896;  
[1]  
SEQUENCE FROM N.A.

**STRAIN=972**;  
MEDLINE=21848401; PubMed=11859360;  
Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A., Brooks J., Peat N., Hayles J., Baker S., Basham D., Bowman S., Skroks K., Brown D., Brown S., Chillingworth T., Churcher C.M., Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A., Gentile S., Goble A., Hamlin N., Harris B., Hidalgo J., Hodgson G., Holroyd S., Hornaby T., Howarth S., Huckle E.J., Hunt S., Jagels K., James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C., Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Squares K., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Volkart G., Aert R., Robben J., Grymonprez B., Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H., Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B., Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M., Lucas M., Rochet M., Gaillardin C., Tallada V.A., Gazzon A., Thode G., Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L., Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J., Shpakovskiy G.V., Ussery D., Barrell B.G., Nurse P.;  
"The genome sequence of Schizosaccharomyces pombe."

**RL** Nature 415:871-880(2002).  
-----  
**CC** This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).  
-----  
**DR** EMBL; AL096788; CAB46672.1; --  
**SQ** SEQUENCE 238 AA; 26479 MW; 58095AA8CD708180 CRC64;

Query Match 62.5%; Score 40; DB 1; Length 238;  
Best Local Similarity 41.7%; Pred. No. 5.5;  
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

**QY** 1 CFQLKKMKVKVR 12  
| : : | | | |  
**Db** 155 CYELQQNSKKIK 166

**RESULT 7**  
**SECY-CVACA** STANDARD; PRT; 410 AA.  
**ID** P46249; Q9MD55;  
**AC** 01-NOV-1995 (Rel. 32, Created)  
**DT** 16-OCT-2001 (Rel. 40, Last sequence update)  
**DT** 15-JUN-2002 (Rel. 41, Last annotation update)  
**DE** Preprotein translocase secy subunit.  
**GN** SECY.  
**OS** Cyanidium caldarium.  
**OG** Chloroplast.  
**OC** Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;  
**OC** Cyanidium.  
**OX** NCBI\_TaxID=2771;  
**RN** [1]  
**RP** SEQUENCE FROM N.A.  
**RC** STRAIN=RK-1;  
**RC** MEDLINE=97134960; PubMed=8980520;  
**RX** Vogel H., Fischer S., Valentin K.-U.;  
**RA** "A model for the evolution of the plastid sec apparatus inferred from secy gene phylogeny."  
**RT** Plant Mol. Biol. 32:685-692(1996).  
**RL** [2]  
**RP** SEQUENCE FROM N.A.  
**RP** STRAIN=RK-1;  
**RC** MEDLINE=20496959; PubMed=11040290;  
**RX** Gloeckner G., Rosenthal A., Valentin K.-U.;  
**RA** "The structure and gene repertoire of an ancient red algal plastid genome."  
**RT** J. Mol. Evol. 51:382-390(2000).  
**RL** CC  
**CC** -!- FUNCTION: INVOLVED IN PROTEIN EXPORT. PROBABLY INTERACTS WITH OTHER PROTEINS TO ALLOW THE TRANSLLOCATION OF PROTEINS ACROSS THE CHLOROPLAST ENDOPLASMIC RETICULUM (CER) MEMBRANES.  
**CC** -!- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast.  
**CC** -!- SIMILARITY: BELONGS TO THE SECY/SEC61-ALPHA FAMILY.  
-----  
**CC** This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).  
-----  
**DR** EMBL; AF022186; AAF12924.1; --  
**DR** InterPro; IPR002208; SecY.  
**DR** Pfam; PF00344; secy; 1.  
**DR** PRINTS; PR00303; SECYTNNLCASE.  
**DR** TIGRFAMs; TIGR00967; 3a0501s007; 1.  
**DR** PROSITE; PS00755; SECY 1; 1.

DR PROSITE, PS00756; SECY 2; 1.  
 KW Protein transport; Transmembrane; Chloroplast; Translocation.  
 FT CONFLICT 149 MISSING (IN REF. 1).  
 SQ SEQUENCE 410 AA; 46242 MW; 7CB0130175BIDF03 CRC64;

Query Match 60.9%; Score 39; DB 1; Length 410;  
 Best Local Similarity 60.0%; Pred. No. 13;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CFQLKKNMK 10  
 ||:|:|:|  
 Db 313 CPELSNNLKK 322

## RESULT 8

ID NOV\_COTJA STANDARD; PRT; 353 AA.  
 AC P42642;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE NOV protein precursor (Nephroblastoma overexpressed gene protein).  
 GN NOV.  
 OS Coturnix coturnix japonica (Japanese quail).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 CC Coturnix.  
 DX NCBI\_TaxID=93934;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA Weiskirchen R., Bister K.;

RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL GROWTH REGULATION (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING

CC PROTEIN FAMILY. CEP-10/CYR61/CTGF/FISP-12/NOV PROTEIN SUBFAMILY.

CC -!- SIMILARITY: CONTAINS 1 VWC DOMAIN.

CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; U13063; AAA2128.1; -  
 InterPro; IPR000359; Cys\_knot.  
 InterPro; IPR000867; Insl\_gro\_fac\_pr.  
 InterPro; IPR000884; TSP1.  
 InterPro; IPR001007; VWF\_C.  
 Pfam; PF00007; Cys\_knot\_1.  
 Pfam; PF00090; tsp\_1.  
 Pfam; PF00093; wvc\_1.  
 Pfam; PF00219; IGFBP\_1.  
 SMART; SM00041; CT; 1.  
 SMART; SM00121; IB; 1.  
 SMART; SM00209; TSP1; 1.  
 SMART; SM00214; VWC; 1.  
 PROSITE; PS00222; IGF\_BINDING; 1.  
 PROSITE; PS01185; CTCK\_1; 1.  
 PROSITE; PS01225; CTCK\_2; 1.  
 PROSITE; PS01208; VWF\_C; 1.  
 PROSITE; PS01208; VWF\_C; 1.  
 T SIGNAL 1 26  
 T CHAIN 27 353  
 T DOMAIN 106 172  
 T DOMAIN 260 334  
 T CTCK.  
 T BY SIMILARITY.  
 T DISULFID 260 297  
 T BY SIMILARITY.  
 T DISULFID 277 311  
 T BY SIMILARITY.  
 T DISULFID 288 327  
 T BY SIMILARITY.  
 T DISULFID 291 329

FT DISULFID 296 333 BY SIMILARITY.  
 FT CARBOHYD 276 276 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 353 AA; 38667 MW; 7179F8533882E89 CRC64;

Query Match 59.4%; Score 38; DB 1; Length 353;  
 Best Local Similarity 58.3%; Pred. No. 17;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFQLKKNMKYR 12  
 ||:|:|:|  
 Db 260 CIRTKKMKAVR 271

## RESULT 9

ID ADDB\_BACSU STANDARD; PRT; 1166 AA.  
 AC P23477;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE ATP-dependent nuclease subunit B.  
 GN ADDB.  
 OS Bacillus subtilis.  
 CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 CC NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OG1;  
 RX MEDLINE=91267926; PubMed=1646786;  
 RA Kocistira J., Venema G.;

RT "Cloning, sequencing, and expression of Bacillus subtilis genes involved in ATP-dependent nuclease synthesis.";

RL J. Bacteriol. 173:3644-3655(1991).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RA Noback M.A., Terpstra P., Holsappel S., Venema G., Bron S.;

RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RL MEDLINE=98044033; PubMed=9384377;

Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S., Borris R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T., Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N., Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G., Guiseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C., Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mabel C., Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetelle D., Porellik S., Prescott A.M., Presecan E., Pujic P., Furnelle B., Rapoport G., Rey M., Reynolds S., Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadate Y., Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F., Sekiguchi J., Sekowska A., Sertor S.J., Sertor P., Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;

RT "The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.";

RL Nature 390:249-256(1997).

1- FUNCTION: THE ENZYME COMPLEX MAY HAVE A WIDE VARIETY OF CATALYTIC  
 ACTIVITIES INCLUDING ATP-DEPENDENT EXONUCLEASE, ATP-STIMULATED  
 ENDONUCLEASE, ATP-DEPENDENT HELICASE AND DNA-DEPENDENT ATPASE  
 ACTIVITIES.  
 2- SUBUNIT: THE B-SUBUNIT IS ATP-DEPENDENT NUCLEASE COMPLEX IS FORMED  
 BY ONE B SUBUNIT IN CONJUNCTION WITH ONE A SUBUNIT.  
 3- This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 4- EMBL: M63489; AAB22200.1; --  
 5- EMBL: Y14081; CAA74481.1; --  
 6- EMBL: Z99109; CAB12902.1; --  
 7- PIR: A39432; A39432.  
 8- Subtilist; BG10465; addB.  
 9- InterPro: IPR000212; UvrD-helicase.  
 10- Pfam: PF00580; UvrD-helicase; 1.  
 11- Hydrolase; Nuclease; Exonuclease; Endonuclease; Helicase; ATP-binding;  
 12- DNA repair; Complete proteome.  
 13- NP BIND 1 22 ATP (POTENTIAL).  
 14- Q SEQUENCE 1166 AA; 134631 MW; 1A57BBE81A08AB4A CRC64;  
 15- Query Match 59.4%; Score 38; DB 1; Length 1166;  
 16- Best Local Similarity 63.6%; Pred. No. 51;  
 17- Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 18- Y 2 FOLKNNKKVR 12  
 19- |||:|:|:|:  
 20- b 486 FOLKNNKKAK 496  
 21- RESULT 10  
 22- 125 MYCCA STANDARD; PRT; 267 AA.  
 23- C P53661;  
 24- T 01-OCT-1996 (Rel. 34, Created)  
 25- T 01-OCT-1996 (Rel. 34, Last sequence update)  
 26- T 01-OCT-1996 (Rel. 34, Last annotation update)  
 27- E Hypothetical protein (Fragment).  
 28- S Mycoplasma capricolum.  
 29- C Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;  
 30- C Entomoplasmataceae.  
 31- X NCBI\_TaxID=2095;  
 32- N [1]  
 33- P SEQUENCE FROM N.A.  
 34- C STRAIN=ATCC 27343 / KID;  
 35- X MEDLINE=96059641; PubMed=7476192;  
 36- A Bork P., Ouzounis C., Casari G., Schneider R., Sander C.,  
 37- Dolan M., Gilbert W., Gillevet P.M.;  
 38- "Exploring the Mycoplasma capricolum genome: a minimal cell reveals  
 39- its physiology".  
 40- L Mol. Microbiol. 16:955-967(1995).  
 41- 1- SIMILARITY: BELONGS TO THE COF/YBHA/YIDA/YIGL (E. COLI) / YCSE/YXEH  
 42- (B. SUBTILIS) FAMILY.  
 43- This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 44- EMBL: Z33006; CAA83689.1; --  
 45- InterPro: IPR001454; Hlgase/hydrolase.  
 46- InterPro: IPR000150; Hypothet\_cof.  
 47- Pfam: PF00702; Hydrolase; 1.  
 48- PROSITE; PS01228; COF\_1; 1.  
 49- Query Match 57.8%; Score 37; DB 1; Length 510;  
 50- Best Local Similarity 77.8%; Pred. No. 36;  
 51- SEQUENCE 510 AA; 58002 MW; A95D675FEA63C668 CRC64;  
 52- Y504 BORBU STANDARD; PRT; 510 AA.  
 53- AC OS1457;  
 54- DT 16-OCT-2001 (Rel. 40, Created)  
 55- DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 56- DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 57- DE Hypothetical protein BB0504.  
 58- GN BB0504.  
 59- OS Borrelia burgdorferi (Lyme disease spirochete).  
 60- NC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 61- CX NCBI\_TaxID=139;  
 62- RN [1]  
 63- RP SEQUENCE FROM N.A.  
 64- RC STRAIN=ATCC 35210 / B31;  
 65- RX MEDLINE=98065943; PubMed=9403685;  
 66- RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,  
 67- Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,  
 68- Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D., Hanson M.,  
 69- Peterson J., Kierulff A.R., Quackenbush J., Salzberg S., Hanson M.,  
 70- van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman C.,  
 71- Utterback T., Wathley L., McDonald L., Artiach P., Bowman J.,  
 72- Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
 73- Smith H.O., Venter J.C.;  
 74- "Genomic sequence of a Lyme disease spirochete, Borrelia  
 75- burgdorferi".  
 76- RT Nature 390:580-586(1997).  
 77- RL 1- SIMILARITY: BELONGS TO THE UPF0144 FAMILY.  
 78- CC 1- SIMILARITY: CONTAINS 1 HD DOMAIN.  
 79- CC 1- SIMILARITY: CONTAINS 1 KH DOMAIN.  
 80- CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 81- between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 82- the European Bioinformatics Institute. There are no restrictions on its  
 83- use by non-profit institutions as long as its content is in no way  
 84- modified and this statement is not removed. Usage by and for commercial  
 85- entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 86- or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 87- EMBL: AE001153; AAC66875.1; --  
 88- TIGR; BB0504; --  
 89- InterPro: IPR002819; HD.  
 90- InterPro: IPR004087; KH dom.  
 91- InterPro: IPR004088; KH type 1.  
 92- InterPro: IPR003607; ME Pplase\_Hdc.  
 93- Pfam: PF00013; KH-domain; 1.  
 94- Pfam: PF01966; HD; 1.  
 95- SMART; SM00471; Hdc; 1.  
 96- SMART; SM00322; KH; 1.  
 97- TIGRFAMs; TIGR00277; HDIG; 1.  
 98- DR PROSITE; PS50084; KH TYPE 1; 1.  
 99- KW Hypothetical protein; Transmembrane; RNA-binding; Complete proteome.  
 100- FT TRANSMEM 2 22 POTENTIAL.  
 101- FT DOMAIN 197 265 KH.  
 102- FT DOMAIN 324 419 HD.  
 103- SQ SEQUENCE 510 AA; 58002 MW; A95D675FEA63C668 CRC64;  
 104- Query Match 57.8%; Score 37; DB 1; Length 510;  
 105- Best Local Similarity 77.8%; Pred. No. 36;  
 106- SEQUENCE 510 AA; 58002 MW; A95D675FEA63C668 CRC64;  
 107- Y504 BORBU STANDARD; PRT; 510 AA.  
 108- AC OS1457;  
 109- DT 16-OCT-2001 (Rel. 40, Created)  
 110- DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 111- DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 112- DE Hypothetical protein BB0504.  
 113- GN BB0504.  
 114- OS Borrelia burgdorferi (Lyme disease spirochete).  
 115- NC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 116- CX NCBI\_TaxID=139;  
 117- RN [1]  
 118- RP SEQUENCE FROM N.A.  
 119- RC STRAIN=ATCC 35210 / B31;  
 120- RX MEDLINE=98065943; PubMed=9403685;  
 121- RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,  
 122- Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,  
 123- Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D., Hanson M.,  
 124- Peterson J., Kierulff A.R., Quackenbush J., Salzberg S., Hanson M.,  
 125- van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman C.,  
 126- Utterback T., Wathley L., McDonald L., Artiach P., Bowman J.,  
 127- Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
 128- Smith H.O., Venter J.C.;  
 129- "Genomic sequence of a Lyme disease spirochete, Borrelia  
 130- burgdorferi".  
 131- RT Nature 390:580-586(1997).  
 132- RL 1- SIMILARITY: BELONGS TO THE UPF0144 FAMILY.  
 133- CC 1- SIMILARITY: CONTAINS 1 HD DOMAIN.  
 134- CC 1- SIMILARITY: CONTAINS 1 KH DOMAIN.  
 135- CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 136- between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 137- the European Bioinformatics Institute. There are no restrictions on its  
 138- use by non-profit institutions as long as its content is in no way  
 139- modified and this statement is not removed. Usage by and for commercial  
 140- entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 141- or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 142- EMBL: AE001153; AAC66875.1; --  
 143- TIGR; BB0504; --  
 144- InterPro: IPR002819; HD.  
 145- InterPro: IPR004087; KH dom.  
 146- InterPro: IPR004088; KH type 1.  
 147- InterPro: IPR003607; ME Pplase\_Hdc.  
 148- Pfam: PF00013; KH-domain; 1.  
 149- Pfam: PF01966; HD; 1.  
 150- SMART; SM00471; Hdc; 1.  
 151- SMART; SM00322; KH; 1.  
 152- TIGRFAMs; TIGR00277; HDIG; 1.  
 153- DR PROSITE; PS50084; KH TYPE 1; 1.  
 154- KW Hypothetical protein; Transmembrane; RNA-binding; Complete proteome.  
 155- FT TRANSMEM 2 22 POTENTIAL.  
 156- FT DOMAIN 197 265 KH.  
 157- FT DOMAIN 324 419 HD.  
 158- SQ SEQUENCE 510 AA; 58002 MW; A95D675FEA63C668 CRC64;  
 159- Query Match 57.8%; Score 37; DB 1; Length 510;  
 160- Best Local Similarity 77.8%; Pred. No. 36;  
 161- SEQUENCE 510 AA; 58002 MW; A95D675FEA63C668 CRC64;  
 162- Y504 BORBU STANDARD; PRT; 510 AA.  
 163- AC OS1457;  
 164- DT 16-OCT-2001 (Rel. 40, Created)  
 165- DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 166- DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 167- DE Hypothetical protein BB0504.  
 168- GN BB0504.  
 169- OS Borrelia burgdorferi (Lyme disease spirochete).  
 170- NC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 171- CX NCBI\_TaxID=139;  
 172- RN [1]  
 173- RP SEQUENCE FROM N.A.  
 174- RC STRAIN=ATCC 35210 / B31;  
 175- RX MEDLINE=98065943; PubMed=9403685;  
 176- RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,  
 177- Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,  
 178- Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D., Hanson M.,  
 179- Peterson J., Kierulff A.R., Quackenbush J., Salzberg S., Hanson M.,  
 180- van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman C.,  
 181- Utterback T., Wathley L., McDonald L., Artiach P., Bowman J.,  
 182- Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
 183- Smith H.O., Venter J.C.;  
 184- "Genomic sequence of a Lyme disease spirochete, Borrelia  
 185- burgdorferi".  
 186- RT Nature 390:580-586(1997).  
 187- RL 1- SIMILARITY: BELONGS TO THE UPF0144 FAMILY.  
 188- CC 1- SIMILARITY: CONTAINS 1 HD DOMAIN.  
 189- CC 1- SIMILARITY: CONTAINS 1 KH DOMAIN.  
 190- CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 191- between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 192- the European Bioinformatics Institute. There are no restrictions on its  
 193- use by non-profit institutions as long as its content is in no way  
 194- modified and this statement is not removed. Usage by and for commercial  
 195- entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 196- or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 197- EMBL: AE001153; AAC66875.1; --  
 198- TIGR; BB0504; --  
 199- InterPro: IPR002819; HD.  
 200- InterPro: IPR004087; KH dom.  
 201- InterPro: IPR004088; KH type 1.  
 202- InterPro: IPR003607; ME Pplase\_Hdc.  
 203- Pfam: PF00013; KH-domain; 1.  
 204- Pfam: PF01966; HD; 1.  
 205- SMART; SM00471; Hdc; 1.  
 206- SMART; SM00322; KH; 1.  
 207- TIGRFAMs; TIGR00277; HDIG; 1.  
 208- DR PROSITE; PS50084; KH TYPE 1; 1.  
 209- KW Hypothetical protein; Transmembrane; RNA-binding; Complete proteome.  
 210- FT TRANSMEM 2 22 POTENTIAL.  
 211- FT DOMAIN 197 265 KH.  
 212- FT DOMAIN 324 419 HD.  
 213- SQ SEQUENCE 510 AA; 58002 MW; A95D675FEA63C668 CRC64;  
 214- Query Match 57.8%; Score 37; DB 1; Length 510;  
 215- Best Local Similarity 77.8%; Pred. No. 36;  
 216- SEQUENCE 510 AA; 58002 MW; A95D675FEA63C668 CRC64;  
 217- Y504 BORBU STANDARD; PRT; 510 AA.  
 218- AC OS1457;  
 219- DT 16-OCT-2001 (Rel. 40, Created)  
 220- DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 221- DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 222- DE Hypothetical protein BB0504.  
 223- GN BB0504.  
 224- OS Borrelia burgdorferi (Lyme disease spirochete).  
 225- NC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 226- CX NCBI\_TaxID=139;  
 227- RN [1]  
 228- RP SEQUENCE FROM N.A.  
 229- RC STRAIN=ATCC 35210 / B31;  
 230- RX MEDLINE=98065943; PubMed=9403685;  
 231- RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,  
 232- Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,  
 233- Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D., Hanson M.,  
 234- Peterson J., Kierulff A.R., Quackenbush J., Salzberg S., Hanson M.,  
 235- van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman C.,  
 236- Utterback T., Wathley L., McDonald L., Artiach P., Bowman J.,  
 237- Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
 238- Smith H.O., Venter J.C.;  
 239- "Genomic sequence of a Lyme disease spirochete, Borrelia  
 240- burgdorferi".  
 241- RT Nature 390:580-586(1997).  
 242- RL 1- SIMILARITY: BELONGS TO THE UPF0144 FAMILY.  
 243- CC 1- SIMILARITY: CONTAINS 1 HD DOMAIN.  
 244- CC 1- SIMILARITY: CONTAINS 1 KH DOMAIN.  
 245- CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 246- between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 247- the European Bioinformatics Institute. There are no restrictions on its  
 248- use by non-profit institutions as long as its content is in no way  
 249- modified and this statement is not removed. Usage by and for commercial  
 250- entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 251- or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 252- EMBL: AE001153; AAC66875.1; --  
 253- TIGR; BB0504; --  
 254- InterPro: IPR002819; HD.  
 255- InterPro: IPR004087; KH dom.  
 256- InterPro: IPR004088; KH type 1.  
 257- InterPro: IPR003607; ME Pplase\_Hdc.  
 258- Pfam: PF00013; KH-domain; 1.  
 259- Pfam: PF01966; HD; 1.  
 260- SMART; SM00471; Hdc; 1.  
 261- SMART; SM00322; KH; 1.  
 262- TIGRFAMs; TIGR00277; HDIG; 1.  
 263- DR PROSITE; PS50084; KH TYPE 1; 1.  
 264- KW Hypothetical protein; Transmembrane; RNA-binding; Complete proteome.  
 265- FT TRANSMEM 2 22 POTENTIAL.  
 266- FT DOMAIN 197 265 KH.  
 267- FT DOMAIN 324 419 HD.  
 268- SQ SEQUENCE 510 AA; 58002 MW; A95D675FEA63C668 CRC64;  
 269- Query Match 57.8%; Score 37; DB 1; Length 510;  
 270- Best Local Similarity 77.8%; Pred. No. 36;  
 271- SEQUENCE 510 AA; 58002 MW; A95D675FEA63C668 CRC64;  
 272- Y504 BORBU STANDARD; PRT; 510 AA.  
 273- AC OS1457;  
 274- DT 16-OCT-2001 (Rel. 40, Created)  
 275- DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 276- DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 277- DE Hypothetical protein BB0504.  
 278- GN BB0504.  
 279- OS Borrelia burgdorferi (Lyme disease spirochete).  
 280- NC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 281- CX NCBI\_TaxID=139;  
 282- RN [1]  
 283- RP SEQUENCE FROM N.A.  
 284- RC STRAIN=ATCC 35210 / B31;  
 285- RX MEDLINE=98065943; PubMed=9403685;  
 286- RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,  
 287- Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,  
 288- Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D., Hanson M.,  
 289- Peterson J., Kierulff A.R., Quackenbush J., Salzberg S., Hanson M.,  
 290- van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman C.,  
 291- Utterback T., Wathley L., McDonald L., Artiach P., Bowman J.,  
 292- Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
 293- Smith H.O., Venter J.C.;  
 294- "Genomic sequence of a Lyme disease spirochete, Borrelia  
 295- burgdorferi".  
 296- RT Nature 390:580-586(1997).  
 297- RL 1- SIMILARITY: BELONGS TO THE UPF0144 FAMILY.  
 298- CC 1- SIMILARITY: CONTAINS 1 HD DOMAIN.  
 299- CC 1- SIMILARITY: CONTAINS 1 KH DOMAIN.  
 300- CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 301- between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 302- the European Bioinformatics Institute. There are no restrictions on its  
 303- use by non-profit institutions as long as its content is in no way  
 304- modified and this statement is not removed. Usage by and for commercial  
 305- entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 306- or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 307- EMBL: AE001153; AAC66875.1; --  
 308- TIGR; BB0504; --  
 309- InterPro: IPR002819; HD.  
 310- InterPro: IPR004087; KH dom.  
 311- InterPro: IPR004088; KH type 1.  
 312- InterPro: IPR003607; ME Pplase\_Hdc.  
 313- Pfam: PF00013; KH-domain; 1.  
 314- Pfam: PF01966; HD; 1.  
 315- SMART; SM00471; Hdc; 1.  
 316- SMART; SM00322; KH; 1.  
 317- TIGRFAMs; TIGR00277; HDIG; 1.  
 318- DR PROSITE; PS50084; KH TYPE 1; 1.  
 319- KW Hypothetical protein; Transmembrane; RNA-binding; Complete proteome.  
 320- FT TRANSMEM 2 22 POTENTIAL.  
 321- FT DOMAIN 197 265 KH.  
 322- FT DOMAIN 324 419 HD.  
 323- SQ SEQUENCE 510 AA; 58002 MW; A95D675FEA63C668 CRC64;  
 324- Query Match 57.8%; Score 37; DB 1; Length 510;  
 325- Best Local Similarity 77.8%; Pred. No. 36;  
 326- SEQUENCE 510 AA; 58002 MW; A95D675FEA63C668 CRC64;  
 327- Y504 BORBU STANDARD; PRT; 510 AA.  
 328- AC OS1457;  
 329- DT 16-OCT-2001 (Rel. 40, Created)  
 330- DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 331- DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 332- DE Hypothetical protein BB0504.  
 333- GN BB0504.  
 334- OS Borrelia burgdorferi (Lyme disease spirochete).  
 335- NC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 336- CX NCBI\_TaxID=139;  
 337- RN [1]  
 338- RP SEQUENCE FROM N.A.  
 339- RC STRAIN=ATCC 35210 / B31;  
 340- RX MEDLINE=98065943; PubMed=9403685;  
 341- RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,  
 342- Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,  
 343- Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D., Hanson M.,  
 344- Peterson J., Kierulff A.R., Quackenbush J., Salzberg S., Hanson M.,  
 345- van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman C.,  
 346- Utterback T., Wathley L., McDonald L., Artiach P., Bowman J.,  
 347- Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
 348- Smith H.O., Venter J.C.;  
 349- "Genomic sequence of a Lyme disease spirochete, Borrelia  
 350- burgdorferi".  
 351- RT Nature 390:580-586(1997).  
 352- RL 1- SIMILARITY: BELONGS TO THE UPF0144 FAMILY.  
 353- CC 1- SIMILARITY: CONTAINS 1 HD DOMAIN.  
 354- CC 1- SIMILARITY: CONTAINS 1 KH DOMAIN.  
 355- CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 356- between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 357- the European Bioinformatics Institute. There are no restrictions on its  
 358- use by non-profit institutions as long as its content is in no way  
 359- modified and this statement is not removed. Usage by and for commercial  
 360- entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 361- or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 362- EMBL: AE001153; AAC66875.1; --  
 363- TIGR; BB0504; --  
 364- InterPro: IPR002819; HD.  
 365- InterPro: IPR004087; KH dom.  
 366- InterPro: IPR004088; KH type 1.  
 367- InterPro: IPR003607; ME Pplase\_Hdc.  
 368- Pfam: PF00013; KH-domain; 1.  
 369- Pfam: PF01966; HD; 1.  
 370- SMART; SM00471; Hdc; 1.  
 371- SMART; SM00322; KH; 1.  
 372- TIGRFAMs; TIGR00277; HDIG; 1.  
 373- DR PROSITE; PS50084; KH TYPE 1; 1.  
 374- KW Hypothetical protein; Transmembrane; RNA-binding; Complete proteome.  
 375- FT TRANSMEM 2 22 POTENTIAL.  
 376- FT DOMAIN 197 265 KH.  
 377- FT DOMAIN 324 419 HD.  
 378- SQ SEQUENCE 510 AA; 58002 MW; A95D675FEA63C668 CRC64;  
 379- Query Match 57.8%; Score 37; DB 1; Length 510;  
 380- Best Local Similarity 77.8%; Pred. No. 36;  
 381- SEQUENCE 510 AA; 58002 MW; A95D675FEA63C668 CRC64;  
 382- Y504 BORBU STANDARD; PRT; 510 AA.  
 383- AC OS1457;  
 384- DT 16-OCT-2001 (Rel. 40, Created)  
 385- DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 386- DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 387- DE Hypothetical protein BB0504.  
 388- GN BB0504.  
 389- OS Borrelia burgdorferi (Lyme disease spirochete).  
 390- NC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 391- CX NCBI\_TaxID=139;  
 392- RN [1]  
 393- RP SEQUENCE FROM N.A.  
 394- RC STRAIN=ATCC 35210 / B31;  
 395- RX MEDLINE=98065943; PubMed=9403685;  
 396- RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,  
 397- Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,  
 398- Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D., Hanson M.,  
 399- Peterson J., Kierulff A.R., Quackenbush J., Salzberg S., Hanson M.,  
 400- van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman C.,  
 401- Utterback T., Wathley L., McDonald L., Artiach P., Bowman J.,  
 402- Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
 403- Smith H.O., Venter J.C.;  
 404- "Genomic sequence of a Lyme disease spirochete, Borrelia  
 405- burgdorferi".  
 406- RT Nature 390:580-586(1997).  
 407- RL 1- SIMILARITY: BELONGS TO THE UPF0144 FAMILY.  
 408- CC 1- SIMILARITY: CONTAINS 1 HD DOMAIN.  
 409- CC 1- SIMILARITY: CONTAINS 1 KH DOMAIN.  
 410- CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 411- between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 412- the European Bioinformatics Institute. There are no restrictions on its  
 413- use by non-profit institutions as long as its content is in no way  
 414- modified and this statement is not removed. Usage by and for commercial  
 415- entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 416- or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 417- EMBL: AE001153; AAC66875.1; --  
 418- TIGR; BB0504; --  
 419- InterPro: IPR002819; HD.  
 420- InterPro: IPR004087; KH dom.  
 421- InterPro: IPR004088; KH type 1.  
 422- InterPro: IPR003607; ME Pplase\_Hdc.  
 423- Pfam: PF00013; KH-domain; 1.  
 424- Pfam: PF01966; HD; 1.  
 425- SMART; SM00471; Hdc; 1.  
 426- SMART; SM00322; KH; 1.  
 427- TIGRFAMs; TIGR00277; HDIG; 1.  
 428- DR PROSITE; PS50084; KH TYPE 1; 1.  
 429- KW Hypothetical protein; Transmembrane; RNA-binding; Complete proteome.  
 430- FT TRANSMEM 2 22 POTENTIAL.  
 431- FT DOMAIN 197 265 KH.  
 432- FT DOMAIN 324 419 HD.  
 433- SQ SEQUENCE 510 AA; 58002 MW; A95D675FEA63C668 CRC64;  
 434- Query Match 57.8%; Score 37; DB 1; Length 510;  
 435- Best Local Similarity 77.8%; Pred. No. 36;  
 436- SEQUENCE 510 AA; 58002 MW; A95D675FEA63C668 CRC64;  
 437- Y504 BORBU STANDARD; PRT; 510 AA.  
 438- AC OS1457;  
 439- DT 16-OCT-2001 (Rel. 40, Created)  
 440- DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 441- DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 442- DE Hypothetical protein BB0504.  
 443- GN BB0504.  
 444- OS Borrelia burgdorferi (Lyme disease spirochete).  
 445- NC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 446- CX NCBI\_TaxID=139;  
 447- RN [1]  
 448- RP SEQUENCE FROM N.A.  
 44

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 LKNNKKVR 12  
 DB 31 LKNNKKVR 39

RESULT 12  
 HKR1\_HUMAN  
 ID HKR1\_HUMAN STANDARD; PRT; 697 AA.  
 AC P10072; Q9UM09;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Krueppel-related zinc finger protein 1 (HKR1 protein) (Fragment).  
 GN HKR1.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Oguri T., Katoh O., Takahashi T., Isobe T., Kuramoto K., Hirata S.,  
 RA Yamakido M., Watanabe H.;  
 RT "The krueppel-type zinc finger family gene, HKR1, is induced in lung  
 RT cancer by exposure to platinum drugs.";  
 RL Submitted (Mar-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 314-537 FROM N.A.  
 RX MEDLINE=8906896; PubMed=2850480;  
 RA Rupert J.M., Kinzler K.W., Wong A.J., Bigner S.H., Kao F.T.,  
 RA Law M.L., Seunazer H.N., O'Brien S.J., Vogelstein B.;  
 RT "The G11-Kruppel family of human genes.";  
 RL Mol. Cell. Biol. 8:3104-3113(1988).  
 CC -!- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTIONAL REGULATION.  
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).  
 CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-  
 CC FINGER PROTEINS.  
 CC -!- SIMILARITY: CONTAINS 1 KRAB DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AB013897; BAA96058.1; -;  
 CC EMBL; M20675; -; NOT\_ANNOTATED\_CDS.  
 CC PIR; C31201; C31201.  
 CC HSP; P08047; 18P2.  
 CC Genew; HGNC:4928; HKR1.  
 CC MIM; 165250;  
 CC InterPro; IPR001909; KRAB.  
 CC InterPro; IPR000822; Znf\_C2H2.  
 CC Pfam; PF00096; zf-C2H2; 13.  
 CC Pfam; PF01352; KRAB; 1.  
 CC SMART; SM00349; KRAB; 1.  
 CC SMART; SM00355; Znf\_C2H2; 13.  
 CC PROSITE; PS00805; ZNF\_C2H2; 1.  
 CC PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 13.  
 CC PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 13.  
 CC Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;  
 CC Nuclear protein; Repeat.  
 CC NON\_TER 1  
 CC DOMAIN 71 142 KRAB.  
 CC ZINC\_FINGERS.  
 CC ZINC\_FINGER 339 695  
 CC C2H2-TYPE.  
 CC ZN\_FING 339 361  
 CC C2H2-TYPE.  
 CC ZN\_FING 395 417  
 CC C2H2-TYPE.  
 CC ZN\_FING 423 445  
 CC C2H2-TYPE.  
 CC ZN\_FING 451 473

FT ZN\_FING 479 501 C2H2-TYPE.  
 FT ZN\_FING 529 C2H2-TYPE.  
 FT ZN\_FING 535 557 C2H2-TYPE.  
 FT ZN\_FING 563 585 C2H2-TYPE.  
 FT ZN\_FING 591 613 C2H2-TYPE.  
 FT ZN\_FING 617 639 C2H2-TYPE.  
 FT ZN\_FING 645 667 C2H2-TYPE.  
 FT ZN\_FING 673 695 C2H2-TYPE.  
 SQ SEQUENCE 697 AA; B574C89FFE55CEB4 CRC64;

Query Match 57.8%; Score 37; DB 1; Length 697;  
 Best Local Similarity 80.0%; Pred. No. 48;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFOLKNNKK 10  
 DB 515 CFSLSKNSLNK 524

RESULT 13  
 TRFL CAMDR  
 ID TRFL CAMDR STANDARD; PRT; 708 AA.  
 AC Q9TUM0; Q9MZS5;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lactotransferrin precursor (lactoferrin).  
 GN LTF.  
 OS Camelus dromedarius (Dromedary) (Arabian camel).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.  
 CX NCBI\_TaxID=9838;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Somali; TISSUE=Lactating mammary gland;  
 RA Kappeler S.R., Ackermann M., Farah Z., Puhan Z.;  
 RT "Sequence analysis of camel (Camelus dromedarius) lactoferrin";  
 RL Int. Dairy J. 9:481-486(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Paramasivam M., Srinivasan A., Singh R., Sahani M.S., Singh T.P.;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH  
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING  
 CC OF AN ANION, USUALLY BICARBONATE (BY SIMILARITY).  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.  
 CC -!- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AJ131674; CAB53387.1; -;  
 CC EMBL; AF165879; AAF82441.1; -;  
 CC HSP; O77811; 1B1X.  
 CC InterPro; IPR001156; Transferrin.  
 CC Pfam; PF00405; transferrin; 2.  
 CC PRINTS; PR00422; TRANSFERRIN.  
 CC SMART; SM00094; TR FER; 2.  
 CC PROSITE; PS00205; TRANSFERRIN\_1; 2.  
 CC PROSITE; PS00206; TRANSFERRIN\_2; 2.  
 CC PROSITE; PS00207; TRANSFERRIN\_3; 2.  
 CC Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;  
 CC Signal.  
 CC SIGNAL 1 19 BY SIMILARITY.  
 CC CHAIN 20 708 LACTOTRANSFERRIN.

```
T REPEAT 20 363 1.
T REPEAT 364 708 2.
T DISULFID 28 64 BY SIMILARITY.
T DISULFID 38 55 BY SIMILARITY.
T DISULFID 134 217 BY SIMILARITY.
T DISULFID 176 192 BY SIMILARITY.
T DISULFID 189 200 BY SIMILARITY.
T DISULFID 250 264 BY SIMILARITY.
T DISULFID 367 399 BY SIMILARITY.
T DISULFID 377 390 BY SIMILARITY.
T DISULFID 424 703 BY SIMILARITY.
T DISULFID 444 666 BY SIMILARITY.
T DISULFID 476 551 BY SIMILARITY.
T DISULFID 500 694 BY SIMILARITY.
T DISULFID 510 524 BY SIMILARITY.
T DISULFID 521 534 BY SIMILARITY.
T DISULFID 592 606 BY SIMILARITY.
T DISULFID 644 649 BY SIMILARITY.
T METAL 79 79 IRON 1 (BY SIMILARITY).
T METAL 111 111 IRON 1 (BY SIMILARITY).
T METAL 211 211 IRON 1 (BY SIMILARITY).
T METAL 272 272 IRON 1 (BY SIMILARITY).
T METAL 414 414 IRON 2 (BY SIMILARITY).
T METAL 452 452 IRON 2 (BY SIMILARITY).
T METAL 545 545 IRON 2 (BY SIMILARITY).
T METAL 614 614 IRON 2 (BY SIMILARITY).
T BINDING 140 140 ANION (BY SIMILARITY).
T BINDING 482 482 ANION (BY SIMILARITY).
T CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
T CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).
T CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).
T CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).
T CONFLICT 261 261 F -> S (IN REF. 2).
T CONFLICT 304 304 G -> A (IN REF. 2).
T CONFLICT 330 330 S -> P (IN REF. 2).
T CONFLICT 492 494 LLS -> PLF (IN REF. 2).
T CONFLICT 506 506 L -> F (IN REF. 2).
T CONFLICT 609 609 A -> P (IN REF. 2).
T CONFLICT 642 642 R -> Q (IN REF. 2).
T SEQUENCE 708 AA; 77211 MW; 0B0C175A0B69D430 CRC64;
Query Match 57.8%; Score 37; DB 1; Length 708;
Best Local Similarity 58.3%; Pred. No. 49;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Y 1 CPOKKNKKVR 12
b 38 CAQWRKKVR 49
RESULT 14
AD5_HUMAN STANDARD; PRT; 784 AA.
P3151;
01-OCT-1993 (Rel. 27, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
3 Vascular endothelial-cadherin precursor (VE-cadherin) (Cadherin-5)
(7B4 antigen) (CD144 antigen).
CDH5.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUE=Endothelial cells;
MEDLINE=95353875; PubMed=7627717;
Breviario F., Cavada L., Corada M., Martin-Padura I., Navarro P.,
Golay J., Introna M., Gullino D., Lampugnani M.G., Dejana E.;
"Functional properties of human vascular endothelial cadherin
(7B4/cadherin-5), an endothelium-specific cadherin.";
Arterioscler. Thromb. Vasc. Biol. 15:1229-1239(1995).
```

```
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=97362755; PubMed=9219219;
RA Ali J., Liao F., Martens E., Muller W.A.;
RT "Vascular endothelial cadherin (VE-cadherin): cloning and role in
RL endothelial cell-cell adhesion.";
RN [3]
RP Microcirculation 4:267-277(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX PubMed=10861224;
RA Shimoyama Y., Tsujimoto G., Kitajima M., Natori M.;
RT "Identification of three human type-II classic cadherins and frequent
RL heterophilic interactions between different subclasses of type-II
RN classic cadherins.";
RX Biochem. J. 349:159-167(2000).
RN [4]
RP SEQUENCE OF 5-784 FROM N.A.
RX TISSUE=Brain;
RN TISSUE=Endothelial cells;
RX MEDLINE=92394977; PubMed=1522121;
RA Lampugnani M.G., Resnati M., Raiteri M., Pigott R., Pisacane A.,
RN Houen G., Ruco L.P., Dejana E.;
RT "A novel endothelial-specific membrane protein is a marker of
RN cell-cell contacts.";
RL J. Cell Biol. 118:1511-1522(1992).
CC -!- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
CC SORTING OF HETEROGENEOUS CELL TYPES. THIS CADHERIN MAY PLAY A
CC IMPORTANT ROLE IN ENDOTHelial CELL BIOLOGY THROUGH CONTROL OF THE
CC COHESION AND ORGANIZATION OF THE INTERCELLULAR JUNCTIONS. IT
CC ASSOCIATES WITH ALPHA-CATENIN FORMING A LINK TO THE CYTOSKELETON.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. FOUND AT CELL-CELL
CC BOUNDARIES AND PROBABLY AT CELL-MATRIX BOUNDARIES.
CC -!- TISSUE SPECIFICITY: ENDOTHelial TISSUES AND BRAIN.
CC -!- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
CC -!- DATABASE: NAME=PROV; NOTE=CD guide CD144 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd144.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X79981; CAA56306.1; --
CC EMBL; U84722; AAB41796.1; --
CC EMBL; AB035304; BAA87418.1; --
CC EMBL; X59796; CAA42468.1; --
CC FIR; S24305; IJHUC5.
CC HSSP; F15116; INCUJ.
CC GlycoSuiteDB; P33151; --
CC Genew; HGNC:1764; CDH5.
CC MIM; 601120; --
CC InterPro; IPR002126; Cadherin.
CC InterPro; IPR000233; Cadherin_C_term.
CC Pfam; PF00028; cadherin; 5.
CC Pfam; PF01049; Cadherin_C_term; 1.
CC PRINTS; PR00205; CADHERIN.
CC SMART; SM00112; CA; 5.
CC PROSITE; PS00232; CADHERIN 1; 3.
CC PROSITE; PS00268; CADHERIN 2; 5.
CC Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
```

KW Signal. 1 25  
FT SIGNAL 26 47  
FT PROPEP 48 784  
FT CHAIN 48 784  
FT DOMAIN 48 784  
FT TRANSMEM 594 620  
FT DOMAIN 621 784  
FT DOMAIN 48 151  
FT DOMAIN 152 258  
FT DOMAIN 259 372  
FT DOMAIN 373 477  
FT DOMAIN 478 593  
FT DOMAIN 736 753  
FT CARBOHYD 61 61  
FT CARBOHYD 112 112  
FT CARBOHYD 157 157  
FT CARBOHYD 362 362  
FT CARBOHYD 442 442  
FT CARBOHYD 523 523  
FT CARBOHYD 535 535  
FT CONFLICT 517 517  
SQ SEQUENCE 784 AA; 87516 MW; F643BFC222A599DE CRC64;

Query Match.  
Best Local Similarity 57.8%; Score 37; DB 1; Length 784;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 FOLKXMKK 10  
Db 379 FOLKXMKK 387

RESULT 15  
YEAS YEAST  
ID YEAS YEAST STANDARD; PRT; 282 AA.  
AC P40003;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DE 01-FEB-1995 (Rel. 31, Last annotation update)  
DE Hypothetical 31.4 kDa protein in GCN4-WBP1 intergenic region.  
ZN YEL005C.  
CS Saccharomyces cerevisiae (Baker's Yeast).  
CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
DX NCBI TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,  
RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,  
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunkle-Smith S.,  
RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,  
RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,  
RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren I., Smith V.,  
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.M.;  
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U18530; AAB64482.1;  
DR SGD; S0000731; VAB2.  
KW Hypothetical protein.  
SQ SEQUENCE 282 AA; 31364 MW; 8F7CA122F91E491A CRC64;

Query Match  
Best Local Similarity 56.2%; Score 36; DB 1; Length 282;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 OLKXMKKV 11  
Db 86 OLKXMKKV 94

Search completed: February 21, 2003, 07:51:44  
Job time : 7.2093 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

M protein - protein search, using sw model

on: February 21, 2003, 07:44:43 ; Search time 21.6744 Seconds  
(without alignments)  
114.078 Million cell updates/sec

file: US-09-743-107b-96  
effect score: 64  
equence: 1 CFQKXNMKKV 12

coring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

searched: 671580 seqs, 206047115 residues

total number of hits satisfying chosen parameters: 671580

inimum DB seq length: 0

aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

atabase :

SPREMBL\_21.\*

1: sp\_archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mhc.\*

8: sp\_organelle.\*

9: sp\_phage.\*

10: sp\_plant.\*

11: sp\_rodent.\*

12: sp\_virus.\*

13: sp\_vertibrate.\*

14: sp\_unclassified.\*

15: sp\_rvirus.\*

16: sp\_bacteriap.\*

17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	43	67.2	711	4	Q8TCD2
2	42	65.6	573	2	Q9S6F6
3	40	62.5	333	5	Q20562
4	39	60.9	38	4	Q3UCY5
5	39	60.9	222	10	Q9LZ65
6	39	60.9	282	16	Q97G23
7	39	60.9	339	10	Q9AW69
8	39	60.9	455	4	Q9BQK9
9	39	60.9	545	4	Q9H117
10	39	60.9	711	4	Q9H116
11	39	60.9	1162	5	Q9S011
12	39	60.9	1553	5	O77385
13	38.5	60.2	1167	10	Q94E36
14	38	59.4	325	10	Q9AYV6
15	38	55.4	2014	5	Q22774
16	37	57.8	108	11	Q9D5E5

17	37	57.8	125	4	Q9H960
18	37	57.8	253	5	Q9VZQ6
19	37	57.8	292	5	Q8SQS0
20	37	57.8	315	5	Q9SSH7
21	37	57.8	335	5	O61888
22	37	57.8	357	13	Q919R0
23	37	57.8	425	4	Q9BSW9
24	37	57.8	546	16	Q98PL7
25	37	57.8	565	16	O8XJU2
26	37	57.8	718	5	Q23447
27	37	57.8	720	6	Q9N003
28	36	56.2	81	16	Q97PW0
29	36	56.2	104	16	Q97HB1
30	36	56.2	124	5	Q9T578
31	36	56.2	129	5	Q9S248
32	36	56.2	129	5	Q8T4U9
33	36	56.2	136	5	O77176
34	36	56.2	167	10	Q9SR76
35	36	56.2	191	16	Q9JQR9
36	36	56.2	192	5	Q8WPY7
37	36	56.2	227	2	Q9X319
38	36	56.2	258	10	O80652
39	36	56.2	260	10	O80651
40	36	56.2	294	3	Q04005
41	36	56.2	295	10	O65376
42	36	56.2	315	16	O8Y311
43	36	56.2	320	2	Q937N3
44	36	56.2	512	13	Q9IAS1
45	36	56.2	538	13	Q9IAS0

## ALIGNMENTS

### RESULT 1

Q8TCD2 PRELIMINARY; PRT; 711 AA.  
AC Q8TCD2;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Lactotransferrin.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PROSTATE;  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC022347; AAH22347.1; -.  
SQ SEQUENCE 711 AA; 78327 MW; 1B9C7EE097C45FAF CRC64;

Query Match 67.2%; Score 43; DB 4; Length 711;  
Best Local Similarity 63.6%; Pred. No. 27;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFQKXNMKKV 11  
||| :|||  
Db 39 CFQWQNRKRV 49

### RESULT 2

Q9S6F6 PRELIMINARY; PRT; 573 AA.  
AC Q9S6F6;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Urease (EC 3.5.1.5) (Urea amidohydrolase).  
GN UREC.  
OS Lactobacillus fermentum.

DC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
 DC Lactobacillaceae; Lactobacillus.

3N [1]\_TaxID=1613;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TK1214;  
 RA Coton E., Visser J.J., Van Vuuren H.J.J.;  
 RT "Urease operon of Lactobacillus fermentum";  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DC -!- CATALYTIC ACTIVITY: UREA + H(2)O = CO(2) + 2 NH(3).  
 CC -!- COFACTOR: EACH CHAIN BINDS TWO NICKEL IONS (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE UREASE FAMILY.  
 CC EMBL; AF120718; AAD22480.1; -;  
 DR HSSP; P41020; 1UBP.  
 DR InterPro; IPR001924; UreaseA.  
 DR Pfam; PF00449; urease; 1.  
 DR PROSITE; PS01120; UREASE\_1; 1.  
 DR PROSITE; PS00145; UREASE\_2; 1.  
 KW Hydrolase; Metal-binding; Nickel.  
 SQ SEQUENCE 573 AA; 61823 MW; 4D325C6C4A21559C CRC64;

Query Match 65.7%; Score 42; DB 2; Length 573;  
 Best Local Similarity 66.7%; Pred. No. 34;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

2Y 1 CFQLKKNKKVR 12  
 [|||||]  
 379 CWQLASKKKVR 390

RESULT 3  
 220562  
 ID Q20562 PRELIMINARY; PRT; 333 AA.  
 AC Q20562;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Hypothetical 38.1 kDa protein.  
 EN F48B6.2.  
 OS Caenorhabditis elegans.  
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 CC Rhabditidae; Peloderinae; Caenorhabditis.  
 CC NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA Geisel C., Bradshaw H.;  
 RT "The sequence of C. elegans cosmid F48B6.";  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA Waterston R.;  
 RT "Direct Submission.";  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U58732; AAB00596.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 333 AA; 38150 MW; B2A5ED57CDBA5C9 CRC64;

Query Match 62.5%; Score 40; DB 5; Length 333;  
 Best Local Similarity 50.0%; Pred. No. 46;  
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 2Y 1 CFQLKKNKKVR 12

Db 198 CFLRKXKVK 209  
 ||:|||||:

RESULT 4  
 Q9UCY5  
 ID Q9UCY5 PRELIMINARY; PRT; 38 AA.  
 AC Q9UCY5;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Lactoferrin homolog (Fragment).  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96081613; PubMed=8551695;  
 RA Sato I.;  
 RT "Characterization of the 84-kDa protein with ABH activity in human  
 RT seminal plasma.";  
 RL Jpn. J. Legal Med. 49:281-293(1995).  
 DR HSSP; P02788; 1BKA.  
 DR InterPro; IPR001156; Transferrin.  
 DR Pfam; PF00405; transferrin; 1.  
 SQ SEQUENCE 38 AA; 4459 MW; 0402F490B5BDBB CRC64;

Query Match 60.9%; Score 39; DB 4; Length 38;  
 Best Local Similarity 63.6%; Pred. No. 11;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 POLKKNKKVR 12  
 ||:|||||:  
 Db 21 FQWQNNKKVR 31

RESULT 5  
 Q9LZ65  
 ID Q9LZ65 PRELIMINARY; PRT; 222 AA.  
 AC Q9LZ65;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Rna binding protein-like (AT5G04600/T32M21\_200) (Putative RNA binding  
 DE protein)  
 GN T32M21\_200 OR AT5G04600.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 CC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Terry N., Ardiles W., Buysshaert C., Dasseville R.,  
 RA De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,  
 RA Villarroel R., Gielens J., Van Montagu M., Bancroft I., Mewes H.W.,  
 RA Rudd S., Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Banh J.,  
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,  
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,  
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,  
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,  
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
 RA Ecker J.R.;



"Arabidopsis ORF clones.";  
 Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 [4]  
 SEQUENCE FROM N.A.  
 Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,  
 Banh J., Bowser L., Carninci P., Chung M.K., Goldsmith A.D.,  
 Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,  
 Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,  
 Nguyen M., Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T.,  
 Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,  
 Yu G., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
 "Arabidopsis cDNA clones.";  
 Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 [5]  
 SEQUENCE FROM N.A.  
 Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,  
 Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,  
 Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,  
 Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,  
 Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,  
 Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,  
 Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,  
 Theologis A.;  
 "Full length cDNA of gene At5g04600 (GI:15238220).";  
 Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 EMBL; AL162875; CAB85566.1; -  
 EMBL; AY055787; AAL06954.1; -  
 EMBL; AF378899; AAK55702.1; -  
 EMBL; AY063822; AAL36178.1; -  
 InterPro: IPR000504; RNA\_rec\_mot.  
 Pfam: PF00076; rrm; 1.  
 SMART; SM00360; RRM; 1.  
 PROSITE; PS0102; RRM; 1.  
 SEQUENCE 222 AA; 25262 MW; BA47EAE6D8B2F153 CRC64;  
 Query Match 60.9%; Score 39; DB 10; Length 222;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 4 LKONKMKV 11  
 |||||  
 9 LKONKMKV 16  
 SULT 6  
 7G23  
 Q97G23 PRELIMINARY; PRT; 282 AA.  
 Q97G23;  
 01-OCT-2001 (TrEMBLrel. 18, Created)  
 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
 Uncharacterized conserved protein.  
 CAC2549.  
 Clostridium acetobutylicum.  
 Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
 Clostridiales; Clostridiaceae; Clostridium.  
 NCBI\_TaxID=1488;  
 [1]  
 SEQUENCE FROM N.A.  
 STRAIN=ATCC 824 / DSM 792 / VKM B-1787;  
 MEDLINE=21359325; PubMed=11466286;  
 Noelling J., Breton G., Omeichenko M.V., Makarova K.S., Zeng Q.,  
 Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,  
 Tarusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,  
 Bennett G.N., Koonin E.V., Smith D.R.;  
 "Genome sequence and comparative analysis of the solvent-producing  
 bacterium Clostridium acetobutylicum.";  
 J. Bacteriol. 193:4823-4838(2001).  
 EMBL; AE007753; AAK80500.1; -  
 Complete proteome.  
 SEQUENCE 282 AA; 31629 MW; E94EE94D502A5730 CRC64;  
 Query Match 60.9%; Score 39; DB 16; Length 282;

Best Local Similarity 58.3%; Pred. No. 60;  
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CFQKKNMKV 12  
 |||||  
 DB 241 CQQLKENISKLR 252  
 RESULT 7  
 Q9AW69 PRELIMINARY; PRT; 339 AA.  
 ID Q9AW69  
 AC Q9AW69  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Hypothetical 41.4 kDa protein.  
 OS Guillardia theta (Cryptomonas phi).  
 OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.  
 OX NCBI\_TaxID=55529;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20087226; PubMed=10618395;  
 RA Zauner S., Fraunholz M., Waati J., Penny S.L., Beaton M.,  
 Cavalier-Smith T., Maier U., Douglas S.;  
 RT "Chloroplast protein and centrosomal genes, a tRNA intron, and odd  
 telomeres in an unusually compact eukaryotic genome, the cryptomonad  
 nucleomorph.";  
 RL Nature 410:1091-1096(2001).  
 DR EMBL; AJ010592; CAC27001.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 339 AA; 41403 MW; B24EE99E30715629 CRC64;  
 Query Match 60.9%; Score 39; DB 10; Length 339;  
 Best Local Similarity 80.0%; Pred. No. 70;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 FQKKNMKV 11  
 |||||  
 DB 179 FNLKONKVKV 188  
 RESULT 8  
 Q9BQK9 PRELIMINARY; PRT; 455 AA.  
 ID Q9BQK9  
 AC Q9BQK9  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE DJ322G13.2.2 (zinc finger protein FLJ21794, isoform 2).  
 GN DJ322G13.2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lovell J.;  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL096677; CAC34610.1; -  
 DR HSP; Q05516; 1CS3.  
 DR InterPro: IPR000210; BTB\_POZ.  
 DR InterPro: IPR000822; ZNF\_C2H2.  
 DR Pfam; PF00651; BTB; 1.  
 DR Pfam; PF00096; zf-C2H2; 5.  
 DR SMART; SM00225; BTB; 1.  
 DR SMART; SM00355; ZNF\_C2H2; 5.

```

R PROSITE; PS00097; BTB; 1.
R PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
R PROSITE; PS00157; ZINC_FINGER_C2H2_2; 5.
R DNA-binding; Metal-binding; Zinc-finger.
Q SEQUENCE 455 AA; 51357 MW; 33099B1B7F7CB757 CRC64;

Query Match
Best Local Similarity 60.9%; Score 39; DB 4; Length 455;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Y 1 CFOLKKQM 8
||| |||
b 125 CFOLKKQM 132

RESULT 9
QH117 PRELIMINARY; PRT; 545 AA.
C QH117;
T 01-MAR-2001 (TRENBLrel. 16, Created)
T 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
T 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
E DJ322G13.2.1 (zinc finger protein FLJ21794, isoform 1).
N DJ322G13.2.
S Homo sapiens (Human).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
X NCBI_TaxID=9606;
P Lovell J.;
A Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

L Submitter (MAR-2001) to the EMBL/GenBank/DBJ databases.
R HSP; P07248; IARF.
R EMBL; AL096677; CAC17422.1; -.
R InterPro; IPR000210; BTB_POZ.
R InterPro; IPR000822; Znf_C2H2.
R Pfam; PF00651; BTB; 1.
R Pfam; PF00096; zf-C2H2; 8.
R ProDom; PD000003; Znf_C2H2; 1.
R SMART; SM00225; BTB; 1.
R SMART; SM00355; Znf_C2H2; 8.
R PROSITE; PS00097; BTB; 1.
R PROSITE; PS00028; ZINC_FINGER_C2H2_1; 7.
R PROSITE; PS00157; ZINC_FINGER_C2H2_2; 8.
R DNA-binding; Metal-binding; Zinc-finger.
W DNA-binding; Metal-binding; Zinc-finger.
Q SEQUENCE 545 AA; 61982 MW; AA9820E319C361E CRC64;

Query Match
Best Local Similarity 60.9%; Score 39; DB 4; Length 545;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Y 1 CFOLKKQM 8
||| |||
b 125 CFOLKKQM 132

RESULT 10
QH116 PRELIMINARY; PRT; 711 AA.
C QH116;
T 01-MAR-2001 (TRENBLrel. 16, Created)
T 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
T 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
E DJ322G13.2.3 (zinc finger protein FLJ21794, isoform 3).
N DJ322G13.2.
S Homo sapiens (Human).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
X NCBI_TaxID=9606;
P Lovell J.;
A Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

```

```

DR EMBL; AL096677; CAC03438.2; -.
DR HSP; P08047; ISP2.
DR InterPro; IPR000210; BTB_POZ.
DR InterPro; IPR000822; Znf_C2H2.
DR Pfam; PF00651; BTB; 1.
DR Pfam; PF00096; zf-C2H2; 10.
DR ProDom; PD000003; Znf_C2H2; 1.
DR SMART; SM00225; BTB; 1.
DR SMART; SM00355; Znf_C2H2; 10.
DR PROSITE; PS00097; BTB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 9.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 10.
KW DNA-binding; Metal-binding; Zinc-finger.
SQ SEQUENCE 711 AA; 80491 MW; 9209B850193BCF1A CRC64;

Query Match
Best Local Similarity 60.9%; Score 39; DB 4; Length 711;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CFOLKKQM 8
||| |||
DB 125 CFOLKKQM 132

RESULT 11
QH5011 PRELIMINARY; PRT; 1162 AA.
ID QH5011;
AC QH5011;
DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE F2966.1 protein.
GN F2966.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1];
RP SEQUENCE FROM N.A.
RA Harris B.R.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [2];
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; Z78543; CAB01753.1; -.
DR HSP; P01003; 3OVO.
DR InterPro; IPR002350; kazal.
DR InterPro; IPR001239; Kazal_inhib.
DR Pfam; PF00050; kazal; 14.
DR PRINTS; PRO0290; KAZALINHTR.
DR SMART; SM00280; KAZAL; 15.
DR PROSITE; PS00282; KAZAL; UNKNOWN_1.
KW Serine protease inhibitor.
SQ SEQUENCE 1162 AA; 130179 MW; 080AEFC277E28CE8 CRC64;

Query Match
Best Local Similarity 60.9%; Score 39; DB 5; Length 1162;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CFOLKKMKVR 12
||| |||
DB 943 CFOLKKMKVTIQ 954

RESULT 12
QH7385 PRELIMINARY; PRT; 1553 AA.
ID QH7385;
AC QH7385;
DT 01-NOV-1998 (TRENBLrel. 08, Created)

```

```

1 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
2 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
3 CDC2-related protein kinase.
4 PFC0755C, MAJ3P6.10.
5 Plasmodium falciparum (isolate 3D7).
6 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
7 NCBI_TaxID=36329;
8 [1]
9 SEQUENCE FROM N.A.
10 STRAIN=3D7; PubMed=10448855;
11 MEDLINE=99376085; PubMed=10448855;
12 Bowman S., Lawson D., Basham D., Brown D., Chillingworth T.,
13 Churcher C.M., Craig A., Davies R.M., Devlin K., Feltham T.,
14 Gentles S., Gilliam R., Hamlin N., Harris D., Holroyd S., Hornsby T.,
15 Horrocks P., Jagals K., Jaseal B., Kyes S., McLean J., Moule S.,
16 Mungall K., Murphy J., Oliver K., Quail M.A., Rajandream M.-A.,
17 Rutter S., Skelton J., Squares R., Squares S., Sulston J.E.,
18 Whitehead S., Woodward J.R., Newbold C., Bartell B.G.;
19 "The complete nucleotide sequence of chromosome 3 of Plasmodium
20 falciparum.";
21 Nature 400:532-538(1999).
22 -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
23 EMBL; Z98551; CAB1141.1; -.
24 InterPro; IPR000719; Euk_PKinase.
25 InterPro; IPR002290; Ser_thr_pkinase.
26 Pfam; PF00069; pkinase; 1.
27 ProDom; PD000001; Euk_PKinase; 2.
28 PROSITE; PS0107; PROTEIN KINASE ATP; 1.
29 PROSITE; PS0011; PROTEIN KINASE DOM; 1.
30 PROSITE; PS00108; PROTEIN KINASE ST; 1.
31 ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
32 SEQUENCE 1553 AA; 182218 MW; E23D3C3416F1E8EF CRC64;
33
34 Query Match 60.9%; Score 39; DB 5; Length 1553;
35 Best Local Similarity 77.8%; Pred. No. 2.5e+02;
36 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
37
38 4 LKONKMKV 12
39 :|||||:
40 39 MKONKMKV 47
41
42 SULT 13
43 4E36
44 Q94E36 PRELIMINARY; PRT; 1167 AA.
45 Q94E36;
46 01-DEC-2001 (TrEMBLrel. 19, Created)
47 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
48 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
49 OSJNB0032H19.13 protein.
50 Oryza sativa (Rice).
51 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
52 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
53 Ehrhartoideae; Oryzeae; Oryza.
54 NCBI_TaxID=4530;
55 [1]
56 SEQUENCE FROM N.A.
57 STRAIN=CV. NIPPONBARE;
58 Sasaki T., Matsumoto T., Yamamoto K.;
59 "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
60 clone:OSJNB0032H19.13";
61 Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
62 EMBL; AP003219; BAB61186.1; -.
63 InterPro; IPR003653; SUMO protease.
64 Pfam; PF02902; Peptidase_C48; 1.
65 SEQUENCE 1167 AA; 131408 MW; 6308AB28A292C1F9 CRC64;
66
67 Query Match 60.2%; Score 38.5; DB 10; Length 1167;
68 Best Local Similarity 72.7%; Pred. No. 2.4e+02;
69 Matches 8; Conservative 2; Mismatches 0; Indels 1; Gaps 1;
70
71 1 CFQLKKNMKV 11
72 :|||||:
73
74 Db 1043 CFQLKKNMKV 1052
75
76 RESULT 14
77 Q9AVY6 PRELIMINARY; PRT; 325 AA.
78 AC Q9AVY6;
79 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
80 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
81 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
82 DE Hypothetical 39.4 kDa protein.
83 OS Guillardia theta (Cryptomonas phi).
84 Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
85 NCBI_TaxID=55529;
86 [1]
87 SEQUENCE FROM N.A.
88 MEDLINE=20087226; PubMed=10618395;
89 RA Zauner S., Fraunholz M., Wastl J., Penny S.L., Beaton M.,
90 Cavalier-Smith T., Maier U., Douglas S.;
91 "Chloroplast protein and centosomal genes, a tRNA intron, and odd
92 telomeres in an unusually compact eukaryotic genome, the cryptomonad
93 nucleomorph.";
94 Proc. Natl. Acad. Sci. U.S.A. 97:200-205(2000).
95 [2]
96 SEQUENCE FROM N.A.
97 MEDLINE=21223349; PubMed=11323671;
98 RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.,
99 Wu X., Reith M., Cavalier-Smith T., Maier U.;
100 "The highly reduced genome of an enslaved algal nucleus.";
101 Nature 410:1091-1096(2001).
102 RL Nature 410:1091-1096(2001).
103 DR EMBL; AJ010592; CAC27085.1; -.
104 KW Hypothetical protein.
105 SEQUENCE 325 AA; 39399 MW; 27D35E07CD82F083 CRC64;
106
107 Query Match 59.4%; Score 38; DB 10; Length 325;
108 Best Local Similarity 54.5%; Pred. No. 1e+02;
109 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
110
111 QY 1 CFQLKKNMKV 11
112 :|||||:
113
114 Db 130 CVLKKNKIKI 140
115
116 RESULT 15
117 Q22774 PRELIMINARY; PRT; 2014 AA.
118 AC Q22774; O01703; O01984;
119 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
120 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
121 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
122 DE Hypothetical 221.3 kDa protein T25C12.3 in chromosome X.
123 T25C12.3.
124 OS Caenorhabditis elegans.
125 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
126 Rhabditidae; Peloderinae; Caenorhabditis.
127 NCBI_TaxID=6239;
128 [1]
129 SEQUENCE FROM N.A.
130 STRAIN=BRISTOL N2;
131 Karslaw J.;
132 Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
133 -!- SIMILARITY: WEAK, TO C.ELEGANS ZK1193.2 AND TO THE C-TERMINAL OF
134 COLLAGENS.
135 EMBL; Z66566; CAA91487.1; -.
136 EMBL; Z78543; CAA91487.1; JOINED.
137 EMBL; Z93779; CAA91487.1; JOINED.
138 EMBL; Z93779; CAB07849.1; -.
139 EMBL; Z78543; CAB07849.1; JOINED.
140 EMBL; Z66566; CAB07849.1; JOINED.
141 EMBL; Z78543; CAB01757.1; -.
142 EMBL; Z93779; CAB01757.1; JOINED.
143 EMBL; Z66566; CAB01757.1; JOINED.
144 WormPep; T25C12.3; CE18966.

```

```

>R InterPro; IPR000561; EGF-like.
>R InterPro; IPR01304; Lectin_C.
>R InterPro; IPR02035; VWF_A.
>R Pfam; PF00059; Lectin_C; 1.
>R SMART; SM00034; CLECT; 1.
>R SMART; SM00181; EGF; 3.
>R SMART; SM00327; VWA; 1.
>R PROSITE; PS00041; C TYPE LECTIN 2; 1.
>R PROSITE; PS00022; EGF 1; UNKNOWN_5.
>R PROSITE; PS01186; EGF 2; 3.
>R PROSITE; PS02034; VWF_A; 1.
>R Hypothetical protein.
>Q SEQUENCE 2014 AA; 221261 MW; 65E3BC76440C73BB CRC64;

Query Match          59.4%; Score 38; DB 5; Length 2014;
Best Local Similarity 87.5%; Pred. No. 4.7e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

>Y 2 FOLKKNMK 9
>b 1410 FOLKKNLK 1417

```

Search completed: February 21, 2003, 08:00:50  
 Job time : 23.6744 secs